Modelling the Causes and Consequences of Paranormal Belief and Experience

Anthony Roy Lawrence

Doctor of Philosophy

University of Edinburgh

1998

Dedication

To Helen, Mum and Dad

Acknowledgements

I would like to thank everyone who helped me during my Ph.D. In particular I would like to thank Professor Robert Morris for his intellectual guidance and example, my friends and colleagues Dr. Chris Roe and Dr. Carl Williams, Dr. Paul Stevens, Zach McDermott, and Simon Sherwood for general support. I'd like to thank Dr. Caroline Watt for taking the time to speak to a lonely seventeen year old conference attendant over tea (she was my first point of contact with the Koestler Chair). I'd really like to thank all my survey participants for having the courage and patience to fill out my questionnaires. I'd like to thank mum and dad full stop, Helen for everything, but especially reminding me how much of my thesis I had to do, Jamie for alcohol soaked nights regaining my lost youth, and Osiris and Nessie (our cats) for being the cause of most of *my* domestic trauma. I would also like to thank my Auntie Doris (who isn't psychic) for part financing my first year of Ph.D.

The research in this thesis was largely funded by a grant from the Economic and Social Research Council, No. R00429324200 (for the last three years), and the Society for Psychical Research (for the first year).

Declaration

Whilst all of the intellectual and analytical work in this thesis is wholly mine, I was assisted in data collection for chapter 5 by Dr. Chris Roe and Dr. Carl Williams (part I) and Pio De Cicco (part II). In chapter 6 I was assisted in data collection by Claire Edwards, Nicholas Barraclough, Sarah Church, and Francesca Hetherington. The following papers have been published as a direct product of research presented in this thesis, and may be found in the appendices (permission for reproduction obtained from the editors of the Journal of Parapsychology, and Personality and Individual Differences);

Lawrence, T.R. (1993) Gathering in the sheep and goats. A Meta-analysis of forced choice sheep-goat ESP studies, 1947-1993. *Proceedings of Presented Papers: The 36th Annual Convention of the Parapsychological Association,* 75-86. (Chapter 10).

Lawrence, T.R. (1995a). How many factors of paranormal belief are there?: A critique of the PBS. *Journal of Parapsychology,* 58, 3-25.(Chapter 3).

Lawrence, T.R. (1995b). Moving on from the Paranormal Belief Scale: A Final Reply to Tobacyk. *Journal of Parapsychology*, 59, 131-140. (Chapter 4).

Lawrence, T.R., Edwards, C., Barraclough, N., Church, S. & Hetherington, F. (1995) Modelling childhood causes of paranormal belief and experience: Childhood trauma and childhood fantasy. *Personality and Individual Differences*, 19, 209-215. (Chapter 6).

Lawrence, T.R., Roe, C.A. and Williams, C. (1997). Confirming the factor structure of the paranormal beliefs scale: Big orthogonal seven or oblique five? *Journal of Parapsychology*, 61, 13-31. (Chapter 5).

Signed:

TABLE OF CONTENTS

Abstract

In this thesis I attempt to develop a basic model of the causes and consequences of paranormal belief and experience, drawing upon the seminal work of Irwin (1992). From an examination of that literature I conclude that research in to paranormal belief in particular, and paranormal experience also to some extent, suffers from being conducted in a piecemeal, atheoretical, and often methodologically weak way. However, firstly I demonstrate that measures of paranormal belief are in a bad state by providing a critique of the most prevalent measure of paranormal belief, Tobacyk and Milford's (1983) Paranormal Belief Scale. I then commence the main empirical component of my thesis. Covariance structure modelling is identified as an excellent tool for the development of models based on correlational survey data, and three studies are conducted which successively refine the modelling of Irwin's (1992) childhood factors model of paranormal belief development. The first study examines a highly simplified version of Irwin's model looking at the role of childhood trauma and fantasy in the development of paranormal experience and belief. The second study provided a small sample confirmation of the results of the first study, in as exact a manner as is statistically possible, and also permitted the piloting of the survey questionnaire that would form the large scale third study. The third study, based on a random street survey of houses in Edinburgh, permitted the final decisive test of both the simplified childhood factors model outlined by Irwin, and my own extended model of the causes and consequences of paranormal belief. In addition the last study tests my own specified trauma hypothesis developed from an examination of results in the first study. This model is found to fit well the final study data, and suggests that there are indeed two routes to the development of paranormal experience; one through fantasy that depends on loss of control trauma, and one direct route from bereavement related trauma. This is probably one of the most important new findings from the thesis. In examining the consequential import of paranormal belief and experience I examine its role in the maintenance of well being and purpose in life, as mediated through personal and interpersonal locus of control. The evidence gained suggests a weak role for paranormal experience in well being and purpose in life, and almost no role for paranormal belief. However, of greater import is the discovery that women and men differ in the direction of their relationship between paranormal factors and well being and purpose in life. Finally I stop to consider the possibility that paranormal experience may be genuine, in order to address this possibility I examine the literature pertaining to the so called sheep-goat effect in ESP testing. A meta-analysis of 73 sheep-goat studies shows that an apparently genuine but very weak sheep-goat effect does exist, and that we can not therefore lightly dismiss the possibility that genuine paranormal phenomena do occur. In concluding I offer up the model developed for further examination, consider its validity in the light of autobiographical accounts of psychics, and allude to several studies by other researchers which have already taken my own work as their starting point.

Chapter 1

INTRODUCTION

> *"It is wonderful that five thousand years have now elapsed since the creation of the world, and still it is undecided whether or not there has ever been an instance of the spirit of any person appearing after death. All argument is against it; but all belief is for it."*
>
> *Samuel Johnson, related by his friend James Boswell*

The Psychology of Paranormal Experience and Belief

For the last twenty years or so psychologists have been paying ever greater attention to the possible causes and consequences of belief in the paranormal (see Irwin, 1993, for a comprehensive and balanced review of research in this area). One suspects this interest is due largely to the persistence of paranormal beliefs despite the great successes of science over the last one hundred years. At the very least this much is said by sceptics seeking to understand why people should hold such 'irrational' beliefs. Of course, parapsychologists have had an interest in paranormal belief almost since the experimental science of parapsychology began, being most specifically concerned with one particular consequence of paranormal belief- its effect on one's ostensible psychic ability (see Schmeidler, 1943; Palmer 1971, 1972; and most recently Lawrence, 1993).

Research into the ostensible paranormal causes and consequences of belief in the paranormal is in itself important enough to justify our interest in this area of investigation, but there are many other lines of research available to the psychologist that justify taking the topic of belief in the paranormal very seriously. Indeed, psychological research on belief in the paranormal has important consequences for our understanding of these commonly held views and the people who hold them. For example, sceptical researchers have made

some fairly strong claims about the psychological constitution of the paranormal believer. Amongst such claims are that paranormal believers are gullible, socially marginal, deficient in powers of rational thought, prone to misperception (in a variety of ways), and of low self-esteem. All of these claims have been made by sceptics in addition to the much stronger claim that paranormal experiencers (especially psychics) are just plain frauds.

When we stop to include those researchers into paranormal belief who are open to the paranormal, the big picture portrays a mixed bag of parapsychologists and psychologists attempting to understand the nature of paranormal belief and experience with a mixed bag of methods, and a mixed bag of results to match. Thus, despite a considerable body of research into the academically correct area of 'belief in the paranormal', we have gained little coherent insight in to the mind of the paranormal believer, and it is perhaps fairer to say that we have often learnt more about the researchers than about their research. As a consequence, we have failed to shed light on an important aspect of human nature. Not parapsychological enough for parapsychologists, and too parapsychological for psychologists, research into paranormal belief and experience has occupied an intellectual no-man's land between the heterodox and the orthodox. To make matters worse, this is a disparate and disjointed community, occupied not by a coherent body of researchers sharing a common outlook but by the more moderate members of the parapsychological and sceptical community, actively pursuing their interests and agendas, and paying little heed to each other.

Parapsychologists have typically ignored the kind of research that forms the basis of this thesis, and there are probably two reasons for this. Firstly, research of this kind is largely conducted by those who are sceptical of the paranormal and thus the broad thrust of what has been found has been neither flattering nor germane to parapsychologists concerns for paranormal phenomena. Secondly, given that parapsychologists' are mostly concerned with the empirical

evidence for paranormal phenomena they have principally advocated a strong experimental approach to these phenomena, in contrast to their Victorian forefathers, the psychical researchers (see Gauld, 1968) who adopted a more eclectic and idiographic approach. Nevertheless, for whatever reason, parapsychologists have typically avoided survey based research, and one can not escape the feeling that the main reason for this is that both psychical researchers and parapsychologists have viewed such research as being irrelevant to the central task of finding firm direct evidence for paranormal phenomena.

I beg to differ however. Parapsychology has attempted for the past fifty years to obtain firm experimental evidence for psi phenomena in the laboratory. That enterprise, for all the effort given over to it, has yielded mixed results. There is a sense of some progress, in some areas, but in general the field has failed to achieve the kind of success that some might initially have expected. Indeed, a number of parapsychologists both sceptical and pro-paranormal have argued that the experimental paradigm be ditched (e.g. Braude, 1986; Blackmore, 1988; White, 1990) in favour of new approaches and paradigms. I myself have also written and spoken on the problems facing contemporary parapsychology (though those efforts are only of oblique relevance to this thesis, see Lawrence, 1995) and my own feelings lie with those who feel that parapsychologists who try and study psi only in the laboratory are casting their net too narrowly. The situation is akin to a hunter who stalks his quarry paying sole reference to the tracks it leaves. Tracks are important, but prey usually leave a great many other clues as to their whereabouts, and the hunter who ignores these signs builds only half a picture. If parapsychologists build only half a picture of their problem area, then it should come as no surprise if they can tell only half a story and consequently catch their quarry only half the time. In this thesis then, I am concerned with introducing new techniques that will allow parapsychologists to build richer narratives. Indeed, the central theoretical and empirical contribution to be outlined in this thesis will be the introduction and use of covariance structure modelling (CSM) to build

'graphical narratives' referring to the hypothesised causes and consequences of paranormal belief.

Covariance structure modelling is no new technique, though its real utility has only become apparent over the last ten years or so. Indeed, psychologists and sociologists have been building causal models using CSM or structural equation modelling for over twenty years now. But parapsychologists have been particularly slow to exploit the new methodological developments in other disciplines, with the consequence that when they have strayed outside the lab and conducted surveys, the results have rarely received the rigorous, theoretically-framed analysis that they have deserved. But let us turn now to a more detailed examination of their effort, as we review the literature on the psychology of paranormal belief and experience.

In the early phase of my researches I quite naturally planned to review the empirical literature on paranormal belief and experience, as would any intending doctoral student. However, in January of 1993 (some three months into my research) Harvey Irwin, noted Australian parapsychologist, produced a quite excellent, comprehensive and balanced survey of just this literature. Consequently, I shall offer here a summary of that review by way of introduction to the area of research into paranormal belief and experience, before going on to delineate my own area of interest, such as will constitute the empirical core of this thesis.

A Review of Research into the Psychology of Paranormal Belief and Experience

This is not a historical work, and I am not given over to the skills and crafts of the historian. Nevertheless a simple acquaintance with the literature of psychology suggests that the psychology of paranormal belief proper began some twenty years ago, with the work of the psychologist Gustav Jahoda in his small book, *The Psychology of Superstition*, published in 1974. The earliest inklings

of such a psychology however, are to be found in the works of another Gustav; Carl Gustav Jung, who in his work *The Psychological Foundation of Belief in Spirits* wrote (somewhat presciently as we shall see),

> *"The 'haunted house' has not yet become extinct even in the most enlightened and the most intellectual cities, nor has the peasant ceased to believe in the bewitching of his cattle. On the contrary, in this age of materialism - the inevitable consequence of rationalistic enlightenment- there has been a revival of the belief in spirits, but this time on a higher level. It is not a relapse into the darkness of superstition, but an intense scientific interest, a need to direct the searchlight of truth on to the chaos of dubious facts. The names of Crookes, Myers, Wallace, Zollner, and many other eminent men symbolize this rebirth...*
>
> *These men typify the reaction of the human mind against the materialistic view of the world. Looked at from the historical standpoint, it is not at all surprising that they used the belief in spirits as the most effective weapon against the mere truth of the senses, for belief in spirits has the same functional significance also for primitive man. His utter dependence on circumstances and environment, the manifold distresses and tribulations of his life, surrounded by hostile neighbours, dangerous beasts of prey, and often exposed to the pitiless forces of nature; his keen senses, his cupidity, his uncontrolled emotions - all these things bind him to the physical realities, so that he is in constant danger of adopting a purely materialistic attitude and becoming degenerate. His belief in spirits, or rather, his awareness of the spiritual world, pulls him again and again out of that bondage in which his senses would hold him; it forces on him the certainty of a spiritual reality whose laws he must observe as carefully and as guardedly as the laws of his physical environment.'*

Even at this prenatal moment in the history of interest in paranormal belief, Jung was placing the prevalence of such beliefs firmly at the door of humanity's struggle with nature. As we shall see, Jung strikes a vein here which has only recently been rediscovered, and which forms the central empirical and theoretical component of this thesis, but I should not reveal the moral before

the story! So, turning to Irwin's (1993) review, he begins by defining the term paranormal;

> *'The term paranormal refers to hypothesized processes that in principle are physically impossible or outside the realm of human capabilities as presently conceived by conventional scientists (Thalbourne, 1982) (Irwin, 1993, pp.1)'*

As far as definitions of paranormal go, Irwin's definition seems to me to be a very good one. I am most impressed by Irwin's definition, over that of others, by the fact that it stresses the human nature of paranormal phenomena. Unfortunately, as Irwin himself acknowledges, many researchers in this area have often employed definitions of the term paranormal that include such things as UFOs, the Loch Ness monster, and Bigfoot. This all-inclusive definition of paranormal is, I think, all too often used to discredit the more credible paranormal phenomena by associating them with anomalous phenomena of less evidential worth. In any case, I shall have cause to touch upon the issue of defining paranormal in a later chapter.

Having defined what he means by the term paranormal, Irwin goes on to discuss the variety of measures that have been developed (I use that term loosely) to assess paranormal belief, and notes in particular their lack of consistency. This presents considerable problems for the interpretation of results drawn from this literature, as we must pay great attention to just what was being measured as paranormal before we can claim to draw any general findings. As we shall see, a considerable portion of my research came to be taken up with a methodological critique of the most prevalent of the present measures of paranormal belief, the Paranormal Belief Scale (PBS) of Tobacyk and Milford (1983; & Tobacyk, 1988), and a good part of that critique focused on the problems of defining *paranormal.* Nevertheless, we should now turn to a concrete discussion of Irwin's review.

Irwin breaks the main body of his review in to four sections; demographic correlates, other belief systems, cognitive variables, and dimensions of personality. In the first of these, the section on demographic correlates of paranormal belief, Irwin notes that for the most part such research has been informed by the *social marginality hypothesis*. According to the adherents of this view (e.g. Bainbridge, 1978; Wuthnow, 1976) people assume paranormal beliefs because they have a marginal social status. Thus the poor, unemployed, and members of ethnic minorities may all show a higher endorsement of paranormal beliefs. This is presumably because such beliefs are hypothesised to be akin to more traditional notions of religious belief, such as that our lot is predestined, the lord will provide, or that the meek will inherit the earth. In this sense paranormal beliefs (in particular beliefs about magical forces or psychic type powers, and beliefs about an afterlife) may offer the socially marginal person some compensation against their lowly social status.

The social marginality hypothesis offers itself up for a fairly clear empirical examination, and Irwin goes on to examine a number of areas where the hypothesis has been, wittingly or unwittingly, put to the test. Thus with respect to age one might expect there to be some association between levels of paranormal belief and one's age. Reviewing the evidence, Irwin notes that several studies have indeed found a significant negative correlation between age and level of paranormal belief. Thus, Emmons and Sobal (1981, p. 52) report age to be the strongest demographic correlate of paranormal belief (negatively so) in their American sample, with findings in agreement from several other studies also (Randall, 1990; Tobacyk, Pritchett, & Mitchell, 1988). However, the negative correlation does not hold up for all of the beliefs that might be termed paranormal (e.g. UFO beliefs, see Randall, 1990).

A significant problem, which Irwin notes, with all previous studies looking at this relationship is that they have been cross-sectional rather than longitudinal, and thus it is very difficult to know whether the correlations that are found are

due to life-long developmental changes, or generational influences, or both. And Irwin is quite right to note that a negative correlation between belief and age is in contradiction to any viable prediction of the social marginality view. If ageism tends to marginalise the elderly, as is generally agreed, then it is they and not the young who should show the highest levels of paranormal belief. In short, the evidence for a relationship between paranormal belief and age fails at all to support the marginality hypothesis.

Nevertheless, there are many ways in which one might be regarded as socially marginalised, and perhaps the most salient of these (insofar as it affects the largest number of people) is the issue of gender differences. In the West at least women are regarded widely as having occupied a position of social marginality for many centuries. Though the women's movement, through feminism, has lead to greater awareness of the problem, and some change in women's equality in the workplace, one could argue that women still have a socially marginal role in an essentially patriarchal world. If this view can be upheld then it would seem the social marginality hypothesis might have an explanatory role, for women do indeed show consistently higher levels of paranormal belief than men (see Blackmore, 1991, for a review).

Interestingly, research has tended to show that whilst women show stronger paranormal belief for humanly relevant paranormal phenomena (like ESP and PK, especially the former) men show a much stronger endorsement than women for more naturally relevant phenomena, such as UFOs and the Loch Ness Monster. The social marginality hypothesis needs to be able to account for this fact, and no one has offered any particularly compelling explanation of this differentiation that would fit with social marginality. Indeed, the so called gender difference in belief may not be a *gender* difference at all, but a basic *sex* difference having more to do with cognitive differences between men and women. Nevertheless, the evidence for gender/sex differences in paranormal belief broadly falls in with the social marginality hypothesis.

Socio-economic status is potentially an important variable in explaining paranormal belief from the marginality perspective. Though it is a tricky concept to operationally define, it is best thought of as a classically latent variable (i.e. one that is not readily apparent, and which only becomes tangible through acquaintance with a number of its most salient indicators, e.g. occupation, income, class. Given the problems of measurement though, one would expect lower socio-economic status to be negatively correlated with paranormal belief. Unfortunately, few studies have looked at socio-economic status and paranormal belief in anything like a satisfactory way. Sheils and Berg (1977) did not find any relationship between global paranormal belief and socio-economic status of students. There are a number of problems with this study though. Firstly, the students' socio-economic status was measured using their *father's* income, occupation and education. Secondly, given the first problem, the lack of any correlation may simply be due to a restriction of range in socio-economic status, especially with a student sample.

In a study by Emmons and Sobal (1981) unemployment was used as an indicator of social marginality, itself an aspect of socio-economic status. The only aspect of paranormal belief that related in a theoretically meaningful way to unemployment was belief in angels. Belief in ghosts, extraordinary lifeforms, and ESP were all less prevalent and salient in the unemployed, contrary to the marginality view.

Apart from these two studies there are a few studies that have explored fairly specific aspects of belief and socio-economic factors (Lundeen and Culdwell, 1930; Ter Keurst, 1939; Zusne and Jones, 1982; Wuthnow, 1976) but I shall not dwell here on their (generally uninteresting) findings as they used dubious indicators of paranormal belief which we shall have later cause to challenge.

Of the other variables that Irwin examines in relation to the marginality hypothesis, there seems to be essentially no good support for the view that

members of ethnic minorities are prone to paranormal belief (Tobacyk, Miller, Murphy and Mitchell, 1988; Emmons and Sobal, 1981; Murphy and Lester, 1976), although differences between cultural and national groupings do occur (Otis and Kuo, 1984; Tobacyk and Pirtilla-Blackman, 1992; Davies, 1988; Irwin, 1991a; Haraldsson, 1985a). However, *these* difference are much more likely to be due to the positive differences in those cultures, than to the negative differences. Thus, unless one can argue that Australian culture conveys a manifest sense of cultural superiority to its members over that of American culture, then the finding that Australians report generally lower paranormal belief than Americans (Irwin, 1991a) can not be taken to support the social marginality hypothesis.

In summary, there seems to be little good evidence to suggest that paranormal beliefs develop as a response to self-perceived social marginality. Perhaps then the explanation for paranormal beliefs is simply that they stem from, and are part of, a general human perspective on the world, one of several which a person might come to adopt. Zusne and Jones (1982) suggest that belief in the paranormal is basically an aspect of a more general subjectivist worldview. Paranormal believers are simply folks who are characterised by an introspective, philosophical, and imaginative outlook on the world, as opposed to an empirical, analytic and determined outlook. Paranormal belief then is to a general worldview what talkativeness is for extroversion, merely a facet. If this is the case then, so Zusne and Jones argue, paranormal belief should be found to cluster along with other subjectivistic beliefs and behaviours.

Once again, Irwin takes it upon himself to review the evidence for this assertion. He starts, rather blandly some may feel, by suggesting that evidence for the worldview hypothesis can be found by looking at the tendency to which strongly held paranormal beliefs go along with strong involvement in paranormally related activities. On this score, one can cite a whole host of evidence to the effect that people who show strong to moderate belief in the

paranormal also show more interest in media relating to the paranormal (books, films, TV programs, etc.), are more likely to take courses on parapsychology, are prone to interpret anomalous experiences as paranormal, have more paranormal experiences, are more likely to use mind altering drugs, and are more likely to practice as a medium or psychic.

It seems to me that from the evidence Irwin cites very little of it can be taken to lend unique and unequivocal support to the worldview hypothesis alone. Indeed, so much of it should be taken as obvious that it can hardly be brought to bare in any explanation; it is rather the thing to be explained. It does perhaps show, as Irwin would have it, that paranormal beliefs are not simply idle philosophies but that they have real associations with the things that people do. Nevertheless, only the finding that paranormal believers show a greater interest in the production of altered states of consciousness (Roney-Dougal, 1984) can be taken as any kind of non-trivial evidence for the world view hypothesis - the rest could be accepted as the basic evidence upon which any good theory of paranormal belief and its concomitants should depend. But the evidence is not all of this trite form, and Irwin goes on to look at the relationship between religious belief and paranormal belief. Although some notable paranormal belief scales include religious beliefs as paranormal beliefs (see my critique of the Paranormal Belief Scale later for some discussion of this problematic point) they are sufficiently different domains of belief that their relationship, or lack of one, would provide suitable grounds for a test of the worldview hypothesis. Thus, if religious belief and paranormal belief didn't correlate to some extent then the worldview hypothesis would be sorely put to find an explanation.

The evidence here is fairly good, but it is complicated by the rather too multidimensional nature of paranormal belief. Thus, in studies of religiosity most studies show that it correlates positively with paranormal belief. The picture is complicated however by the fact that certain subscales of some

paranormal belief scales do not show positive correlations with paranormal belief. Research in to other areas of belief that covary with paranormal belief have shown that ESP belief tends to correlate with dualist mind/body beliefs (Stanovich, 1989), that paranormal believers are more prone to believe in an after life (r=0.42-0.63, Haraldsson, 1981; r=0.26-0.48, Irwin, 1985a), and that they are more prone to self-reflective or absorbing internal experiences (Irwin, 1985b; Irwin, 1989). Lastly, Zusne and Jones developed a World-view scale to test their hypothesis. They found that believers tended to score highly on the subjective component of the scale.

Although Irwin (1993) feels that diverse evidence generally supports the worldview hypothesis, he notes that the theory leaves many things to be desired. On this I agree, only more so. The worldview hypothesis, as for many hypotheses about paranormal beliefs, seems too trite to be taken seriously. It is far too general, and much of the evidence in support of it can be explained equally well, if not better, by other theories. What is essentially lacking in the worldview hypothesis is any interest in what causes certain people to tend towards a subjective worldview. As Irwin is quite right to point out, what fundamental determinants of subjective worldviews give rise to paranormal belief? This kind of enquiry requires a move away from merely establishing that paranormal beliefs go along with other kinds of beliefs that have a subjective basis, to looking at the basic cognitive and personality characteristics that might support such beliefs. That, after all, is what constitutes the true psychological enquiry. It is this kind of evidence that Irwin surveys next, in his section on the cognitive correlates of paranormal belief.

Turning to the cognitive correlates of paranormal belief one comes directly to the centre ground for research conducted by sceptical psychologists. Though by no means driven by an explicit theory, the greatest component of this kind of research can be seen to fit in to a general framework. Thus, sceptical psychologists argue that given that surveys suggest that paranormal belief is

widespread (a well attested fact), and that evidence of the paranormal is at best weak (a somewhat less well attested opinion), then there is a high likelihood that paranormal believers who claim to base their beliefs on experience, or indeed on any evidence, are not doing so on the basis of real experience, but on some kind of ersatz constructed experience. And finally, so the argument goes, much of this ersatz experience and evidence must be a function of the paranormal believer's particular personal makeup. In short, there must be something about believers that predisposes them to believe in the paranormal. Naturally, questions of this sort have lead researchers to examine potential cognitive and personality variables that might predispose people to paranormal like experiences, or bias their interpretation of second hand evidence of paranormal phenomena.

Because of this emphasis on bias, illusion, gullibility and incredulity in the development of paranormal belief, Irwin (1993) has wryly termed this vein of research *cognitive deficits research.* A review of this specific area of research is also provided by French (1992), himself a moderately sceptical psychologist working within the cognitive deficits domain. Irwin (1993) reviews the evidence for cognitive correlates of paranormal belief on a broad front. He partitions his discussion of these correlates in to four sections; educational attainment, scientific and other specialist education, intelligence and reasoning skills, and creativity and imagination.

For the first of these cognitive correlates, educational attainment, Irwin points out a number of studies that have addressed this issue in a variety of ways. Thus, Messer and Griggs (1989) correlated grade point average in US college students with paranormal belief and found a negative correlation, implying that the relative sceptic was a better student than the relative believer. Tobacyk (1984) reports a similar negative relationship between college student grade point average and superstition and spiritualist beliefs on his PBS. Then again, for somewhat younger high school students Tobacyk, Miller and Jones (1984)

found that grade point average correlated positively with belief in psi and traditional religious belief. Irwin argues that further research is warranted to look at the relationship between paranormal beliefs, grades, and developmental stage.

Whilst not wishing to prejudge the issue, I do not share Irwin's optimism about the import of these studies. Firstly, it may well be the case that traditional religious belief and psi beliefs *are* indicative of a generally well educated mind, whilst superstition and spiritualist beliefs are less positively indicated on this front. Certainly the PBS shows differential relationships between its subscales and a number of traits (see Davies and Kirby, 1985, in particular for evidence of this kind). In short, the above findings may simply reflect genuine differences in the way these PBS subscales relate to educational attainment. Secondly, given the well attested relationship between extroversion and belief in ESP (e.g. Thalbourne, 1981; Thalbourne and Haraldsson, 1980) and extroversion and educational attainment (Kline, 1993) the possibility that the relationships found (especially the negative ones) are the result of the underlying connection with extroversion cannot lightly be dismissed. One simple way to test this would be to calculate the partial correlation of educational attainment and paranormal belief with the variance due to extroversion partialled out. If a very small, or no, correlation was found this would strongly suggest that extroversion is the key variable here.

Nevertheless, Irwin goes on to look at other variables that might index educational attainment. The highest level of education achieved is one other way of assessing educational attainment. Here, Irwin notes that previous research had been quite inconsistent. Tobacyk *et al.* (1984) found that paranormal belief was positively related to attainment according to this index. In a related vein, a number of studies (Emmons and Sobal, 1981; Haraldsson, 1985a, p. 149; Tobacyk *et al.*, 1984) have found that psi belief is positively related to highest level of education, whilst a fairly equal number of others

(Gray, 1987; Haraldsson, 1985a, p. 149, Otis and Alcock, 1982; Pasachoff, Cohen, & Pasachoff, 1970) have found a negative relationship. The same mixed pattern is found for most other subscales of paranormal belief, with only Spiritualist beliefs showing any truly consistent (negative) relation to educational attainment (Otis and Alcock, 1982).

One major problem with this whole body of research is that it lacks resolution. Educational attainment is such a gross index of cognitive functioning that it may be far too broad a variable to aid in the identification of those areas of cognitive functioning that relate strongly to paranormal beliefs. As Irwin (1993) himself points out, the problems with educational attainment and paranormal belief are legion; '...depending on the context of its measurement it could be confounded with age, socio-economic status, generational differences, developmental level, social roles associated with particular vocations, and exposure to forms of social indoctrination other than education' (p. 17). Given this, it may be better to identify aspects of educational attainment that would more readily relate to paranormal belief. In the next section, on scientific and other specialist education Irwin does just this.

On first appearances the more scientific education one has, the less likely one is to believe in the paranormal. A scientific education, so the story goes, is the antidote to superstitious thinking of all kinds (Valentine, 1936). Indeed, Singer and Benassi (1981), with a breathtakingly careless disregard for the subtleties of paranormal beliefs, argue that the quality of science teaching in the US can be indexed against the prevalence of paranormal beliefs. A number of studies show that some or other type of paranormal belief is less prevalent amongst those with a college science education (Happs, 1987; Padgett *et al.*, 1981; Otis and Alcock, 1982; Gray and Mill, 1990; Za'rour, 1972) in comparison to those with a humanities background.

However, although differences do occur, these can hardly be explained easily in terms of basic cognitive functions, for once again educational attainment in terms of the faculty of one's education is hardly a sensitive measure of underlying cognitive function. In addition, it is difficult to tease out the causal direction of any possible relation between these variables. What is needed here is a more specific and precise test that directly addresses the question of the basic underlying determinants of paranormal belief. More basic measures of a range of cognitively relevant variables are needed. IQ (visual and verbal, fluid and crystallised), imagery, creativity, and reasoning ability would form the core variables.

Intelligence as measured by IQ tests of various kinds has been identified as one of the most fundamental determinants of human behaviour and thought (Kline, 1991). If one was searching for a basic cognitive trait that might have some bearing upon the development of paranormal belief then IQ would provide an obvious possibility.

Strangely, practically no research has been conducted to look at the relationship between paranormal belief and IQ, visual, verbal, or otherwise defined. Irwin (1993) himself remarks on this dearth of evidence. Though early studies did look at the relationship between IQ and superstitious beliefs (Emme, 1940; Zapf, 1945; Killen, Wildman, and Wildman, 1974), finding them to be negatively related, very few studies have looked at paranormal belief in other contexts. Indeed, what little research does exist is involved primarily with investigating the claim that paranormal believers can't think straight, measured in terms of their ability to perform well in syllogistic reasoning tasks and the like.

Interestingly, one more recent study that has looked at general paranormal belief and IQ was conducted by Jones *et al.* (1977). In that study they found that IQ correlated positively with belief ($r=0.22$). Smith, Foster and Stovin

(1995) however, found that IQ (as measured by the Advanced Raven's Progressive Matrices) correlated negatively with paranormal belief. Smith's study is interesting in that he specifically tried to test Irwin's (1993) claim that the experimenter's expressed belief in the paranormal would moderate the relationship found between IQ and paranormal belief. Smith et al. (1995) did not confirm this view, and so the IQ/belief relationship, though scantily evidenced, does tend to support the view that sceptics show slightly higher IQs than believers, at least for those components of IQ tested (they have tended to be weighted in favour of verbal IQ measures). Other studies have looked more generally at critical reasoning ability (Alcock and Otis, 1980; Gray and Mill, 1990; Polzella, Popp and Hinsman, 1975; Wierzbicki, 1985; Tobacyk and Milford, 1983). These studies tend to find that believers are less critical in their thinking than sceptics.

Some caution needs to be expressed about this kind of finding however, as Tobacyk and Milford (1983) found that only those who scored highly on traditional religious belief and superstition showed significantly poorer critical thinking. Irwin himself argues that these studies, conducted solely by researchers of a sceptical persuasion, may suffer from a particular artefact. In short, if the researcher impresses his or her scepticism upon the participant too much then they may well force otherwise rational believers to profess a false scepticism, in a way which is clearly reminiscent of the socially desirable response. The result will be that but for the sheep in goats clothing there would in fact be little difference between the sheep and the goats on measures of critical inference. This is indeed an interesting proposition, and might help explain why Irwin (1991a) has found little evidence for critical thinking differences between believers and sceptics (save for strong religious beliefs). Unfortunately for Irwin, the only critical test of this hypothesis has found it wanting (Smith, Foster and Stovin, 1995).

To summarise, though there is practically no good research looking at IQ and belief in the paranormal, some research looking at more general cognitive constructs has shown believers to perform poorly on tests of critical and conditional reasoning. This research itself is piecemeal and highly mixed in quality, and it is very difficult to say whether the results obtained so far represent a body of data of central or peripheral importance in understanding paranormal belief.

The previously discussed research addressing cognitive correlates of paranormal belief has been primarily concerned with cognitive deficits, and primarily it is the believer who has been argued to suffer from them. Nevertheless, other kinds of cognition exist apart from the purely propositional, analytic, and numerical kinds (i.e. the kinds that appeal to staunch ratiocentric sceptics), and a fair amount of research has been conducted to look at these. Irwin's final section on the cognitive correlates of paranormal beliefs deals with creativity and imagination. It should come as no surprise that these more positive aspects of cognitive function have been suggested by more moderate or pro-paranormal researchers to be related to paranormal belief and experience. Thus, both Krippner (1962) and Murphy (1963) have suggested that paranormal phenomena are associated with creativity. Likewise, Moon (1975) found that ESP belief was higher in artists than in non-artists. More directly relevant to the creativity issue, Joesting and Joesting (1969) found a positive correlation between ESP belief and a standard measure of creativity, and Davis, Peterson and Farley (1974) found a similar result for psi belief generally.

The data on creativity provide quite unequivocal support for the view that believers are more creative than sceptics, but they do much more than this. Though Irwin fails to note this, these data accord well with experimental tests of ESP that have examined the influence of similar cognitive variables on actual ESP performance. Thus, Schlitz and Honorton (1992) found a stunning 50%

hit rate for artists at the world renowned Julliard School for the Performing Arts in New York, when they tested them in their Ganzfeld laboratory at Psychophysical Research Laboratories (this result was replicated by Radin, McAlpine and Cunningham 1994). This study has spawned a number of other ESP Ganzfeld studies that report results in a similar vein. Creativity actually seems to be associated with laboratory ESP results. For once a cognitive correlate of paranormal belief and experience in self-report measures is found to hold true for controlled laboratory tests. This kind of result surely warrants more serious consideration given its capacity to switch the focus of psychological research on belief more clearly, and honestly, on to paranormal experience.

A second variable that has been found to correlate positively with paranormal belief and experience is fantasy proneness. Irwin (1990, 1991b) found that fantasy proneness as measured by Myers' (1981) ICMI:C scale correlated positively with paranormal belief (range of $r = 0.32$-0.48 for core components of paranormal belief). More generally Wilson and Barber (1983) found that a group of highly susceptible hypnotic subjects (all female) reported having quite florid fantasy lives and often reported plenty of psychic experiences. Hypnotic susceptibility has been found to correlate with paranormal belief and experience in a number of studies (Haraldsson, 1985b: Nadon, Laurence & Perry, 1987; Wagner and Ratzenberg, 1987). This whole body of results suggests that paranormal beliefs are related to introspective, absorbing, imagery based experiences, and the capacity to entertain these fairly easily. More interestingly still, Irwin goes on to note the relationship between fantasy and childhood trauma, as previously noted by Hilgard (1974), Rhue and Lynn (1987) and Wilson and Barber (1983). For example, Rhue and Lynn (1987) argue that, 'It may be that fantasizers' imaginative involvements represent an autistic retreat into a fantasy world as a result of an inability to cope effectively with reality.' (pp. 327). However, Rhue and Lynn (1987) did not find that high fantasy was strongly associated with psychopathology, although 10-20% of high

fantasizers displayed some symptoms of psychopathology. Logically, so Irwin argues, trauma may have an influence on the development of paranormal beliefs, via the influence of fantasy proneness. Support for just such a notion is found in a number of studies (Hilgard, 1974; Wilson and Barber, 1983; Rhue, Lynn, Henry, Buhk and Boyd, 1990-91, Council and Huff, 1990). To further examine this claim Irwin introduces the last of his four major correlates of paranormal belief - the personality correlates.

Of obvious interest when dealing with personality correlates of paranormal belief are the main components of personality - extroversion, neuroticism and psychoticism (at least if one adopts an Eysenckian enumeration of personality). Firstly, with respect to extroversion most of the evidence does not support a relationship with paranormal belief (r=0.05, Windholz and Diamant, 1974; Lester, Thinschmidt, and Trautman, 1987), though Thalbourne and his colleagues have consistently found a small correlation between extroversion and belief in ESP (Thalbourne, 1981; Thalbourne, Beloff, Delanoy, and Jungkuntz, 1983; Thalbourne and Haraldsson, 1980). Interestingly extroversion has been implicated in the susceptibility to paranormal experience (Eysenck, 1967) and a meta-analysis of the experimental ESP-Extroversion literature appears to show a small but genuine relationship (Honorton, Ferrari and Bem, 1991). There has been practically no research examining the other scales of personality according to the three factor structure that Eysenck favours. However, a study some colleagues and I conducted at St. Andrews University did show that extroversion correlated with Thalbourne and Delin's (1993) Australian Sheep-Goat Scale (r = 0.268) and the revised Paranormal Belief Scale (Tobacyk, 1988) (r=0.185). Neither neuroticism or psychoticism correlated with the ASGS or PBS (although Windholz and Diamant, 1974, reported an r=0.21 with their SOBEP scale). Extroversion alone then seems to be related to paranormal belief and experience.

According to Eysenck (1967) the relation between extroversion and paranormal experience can be explained in terms of the low arousal characteristic of extroverts facilitating the receipt of paranormal information. Strangely, researchers, apart from Sargent (1981), have been slow to follow up this type of research, perhaps because an adequate test of the hypothesis would require access to psychophysiological recording techniques that parapsychologists have neither the money nor the expertise to employ. Irwin, however, chooses to think of the extroversion data from a social perspective, and from that perspective he notes that the evidence for the social dimensions of personality in paranormal believers depict an apparently outgoing individual who derives much of their stimulation from their own inner life, rather than their social relationships with others. Irwin feels this characterisation is almost psychopathic, though it is not at all clear to me the evidence Irwin cites either individually or collectively supports this notion. However, there is more direct evidence bearing upon the psychological adjustment of the paranormal believer. Early studies carried out in the thirties on superstition found negative relationships between superstitiousness and adjustment (Maller and Lundeen, 1934; Ter Keurst, 1939; Zapf, 1945), but superstition is one component of the present day Paranormal Belief Scale that shows little variation in response and very low endorsement - few people are superstitious these days it seems. Thus, Irwin ($r=0.05$; 1991a) and Tobacyk and Milford (1983) both fail to support a link between modern day superstition and poor psychological adjustment.

Oddly, Schumaker (1987) argues that paranormal beliefs should lead to enhanced psychological adjustment, and claims to have support for this from studies he has conducted ($r=+0.39$), but Irwin rightly points out that Schumaker has simply interpreted his measure of adjustment incorrectly - when the correct interpretation is used the relationship that Schumaker claims to have found is the reverse (Irwin, 1991a, reports an $r=+0.19$ for his own replication attempt). It should be noted also here that Schumaker's supporting

arguments are so poor, and his hatred for paranormal belief so fervent, that his argument barely warrants scrutiny at all.

More interesting than the evidence for general psychological adjustment is the evidence for a more specific proto-psychopathology in paranormal believers. Evidence does tend to suggest that paranormal believers score more highly on Eckblad and Chapman's (1983) Magical Ideation Scale than do non-believers (Thalbourne, 1985; Anderson, 1988; Thalbourne, Dunbar and Delin 1989; Williams, 1989; Tobacyk and Wilkinson, 1990; Thalbourne, 1994, r=0.54). Most important here is Thalbourne's (1985) finding that the relationship between magical ideation and paranormal belief holds up even when items directly referring to parapsychological phenomena on the magical ideation scale are removed from the analysis. The magical ideation scale is regarded as an indicator of schizotypy, or schizoid personality disorder, a less pronounced variant of full blown schizophrenia. Interestingly, Thalbourne (1994) finds that the correlation between belief and magical ideation were generally much higher for males over females.

Irwin notes that the studies on magical ideation could be taken to mean that believers are just crazy. He quite rightly cautions against this interpretation, citing two possible objections. Firstly, one could be a paranormal believer in a number of ways, and only some kinds of paranormal believer may show a strong magical ideation profile. Thus, Williams and Irwin (1991) compared members of an Australian psychical research group with a sample of schizotypes on a measure of paranormal belief and schizotypy. Whilst both groups had similar belief scores, the psychical researchers' schizotypy scores were the same as those for normal members of the adult population, whilst of course the schizotypy sufferers' scores were reminiscent of an abnormal population. We may therefore need to identify just what kind of paranormal believer we are dealing with, before we jump to the conclusion that they are all crazy. This is precisely a point I bring out in far greater detail in chapter 4. Our

present conception of paranormal belief is simply too clumsy to allow really incisive research to discover what is going on.

Secondly, clearly magical ideation is a relevant variable for understanding (some class of) paranormal believer, but the simple correlation in itself does not tell us anything other than that the two variables are related. As Irwin (1993) points out, recalling points I have raised earlier, 'That paranormal believers tend to be psychologically maladjusted is more an observation to be explained than a sufficient explanation in itself. Given that paranormal beliefs have some attraction for relatively maladjusted (and other) people, under the psychodynamic functions hypothesis the question still remains as to the *functions* served for these people by paranormal beliefs. It is in such terms that an effective theory of paranormal beliefs should be couched.' (p. 25).

Here at last Irwin takes twenty years of paranormal belief research one step further, by pointing out that it would be nice at this stage to pause and consider what the data so far have been telling us. The fact of the matter is that up until quite recently the level of theorising in paranormal beliefs research was shockingly poor. As we have seen, the research can just be organised around loose themes, cognitive deficits, worldview, social marginality, but none of these attempts has been particularly successful, save perhaps for selected studies within the cognitive deficits vein.

In considering the functions that paranormal beliefs might serve Irwin is quick to note that whilst social and interpersonal factors may well influence the development and expression of paranormal beliefs, it is in the psychodynamic (intrapersonal) domain that paranormal beliefs are most likely to find their root cause. A whole host of sceptical researchers have argued that paranormal beliefs may serve an important role in the development of a sense of security in an uncertain world. Thus, if the world is thought to be chaotic or capricious, this can easily overwhelm one's will to cope with life events, creating the sort

of existential angst that beset William James (Fancher, 1979), the great American psychologist. Paranormal beliefs provide a template for understanding capricious and traumatic events. For example, 'only the good die young' is an expression we may use to console those who have lost a son or daughter, but its consolatory value lies in nothing concrete and real, but merely in the quasi-religious assertion that nature (or god) knows our deeds, and can claim us early if we accrue a positive balance of good ones. The thought consoles us because it plays to our better sentiments (our son was a good lad, we should be proud he led such a worthwhile life, he was too good for us etc.). The role that paranormal beliefs play then is just that of 'only the good die young' abstracted and writ large.

Sceptics have not been slow to aver to this potential raison d'être for paranormal beliefs (Alcock, 1981; Frank, 1977; Marks and Kammann, 1980; Schumaker, 1990; Singer and Benassi, 1981; Zusne and Jones, 1982). Few however, have really sought to test this claim properly. Nevertheless, the work of Blackmore and Troscianko (1985) is pertinent in this respect. Blackmore and Troscianko argued that paranormal beliefs would provide an illusion of control over events which were otherwise controlled by chance alone. To test this they asked study participants to estimate the degree of control they had over a computer coin tossing task. In a repeated measures design participants took part in trials where some control was permissible, whilst in others no control was permissible. They found that believers felt they had more control over the computer task than sceptics felt they did, whether they actually did or not, and irrespective of their actual success in the task. Irwin cites his own study (Irwin, 1992) as additional support for the claim that control is a prime concern for paranormal believers. Irwin (1992) argued that the believer's need for control over the world might be manifested in a higher need for control in interpersonal relationships. To test this he administered Schultz's (1957; 1978) FIRO-B scale to some psychology students. Amongst other things, the FIRO-B claims to measure a person's need for control in their interpersonal

relationships in the form of an *expressed control* score. Irwin, concerned about the potential fakeability of the FIRO-B (in my view unnecessarily), provided duplicate forms of the scale to the respondent and a close friend of theirs, in order to obtain two perspectives on the respondents' need for control. Whilst the expressed control scores did correlated positively with respondent paranormal belief when the associate estimated the respondents' need for control, it did not do so when the respondent provided their own responses. Irwin nevertheless argues that the results support the notion that paranormal belief is in some way tied up with a need for control over the world, and feels that these results point to a psychodynamic source for paranormal beliefs. Strangely, Irwin neglects to mention the findings of another study which bears directly on the control theme, that of Davies and Kirby (1985). Davies and Kirby provided 95 liberal arts college students with Paulhus's (1983) Spheres of Control Scale and Tobacyk and Milford's (1983) Paranormal Belief Scale, in an attempt to measure the extent to which perceived control and paranormal belief were related. In a canonical correlation analysis of the two scales they found that internal control in the interpersonal and personal spheres was associated most strongly with psi belief and witchcraft, whilst external control (control by others) in the personal and socio-political spheres is associated most strongly with traditional religious belief, superstition and spiritualism. This study clearly indicates that the more experientially relevant paranormal beliefs are associated with concerns about the personal, and moreover interpersonal domains. To the extent that the core paranormal beliefs (psi and witchcraft being representative here) correlate with control related measures, they seem to be most strongly related to interpersonal control of just the kind that the expressed control score of the FIRO-B is designed to measure.

Given the relation between paranormal belief and the need for control over the interpersonal domain then it is pertinent to ask why this should be the case. Irwin squarely places the cause of this relationship at the door of its potential relevance to the relationship between childhood trauma and fantasy.

Thus, given that studies (previously discussed) implicate trauma in the development of fantasy proneness, and given that fantasy proneness is a trait that is clearly relevant to paranormal belief, then it is not such a great inferential leap to suppose that control over the interpersonal domain may be a feature of paranormal belief because that belief finds its basis in fantasy proneness and fantasy finds (some of) its basis in childhood trauma. This is at least the hypothesis that Irwin puts forward strongly both in his 1993 review and in his 1992 paper that explicitly addresses this issue.

Irwin's *childhood factors model* (as I choose to call it in this thesis) is a reasonably bold hypothesis, because it is both reasonably specific in its claim, and (a priori) sufficiently novel (even exotic) as to warrant the claim that on first appearances it should be found wanting. Nevertheless, Irwin (1992) did find evidence to suggest that childhood trauma has some role to play in explaining the development of paranormal beliefs. Irwin administered Council and Edwards' Survey of Traumatic Childhood Events (STCE) to a group of university students along with Tobacyk's (1988) revised Paranormal Belief Scale. He found that global paranormal belief correlated significantly and positively with intrafamilial physical abuse during childhood, and that overall trauma correlated positively with paranormal belief, although at a marginal level of significance.

As a result of this study, Irwin (1992) proposed a model of the development of paranormal belief that he felt might act as a stimulus for further research. This model is represented exactly as Irwin depicted it in his 1993 review, in figure 1.

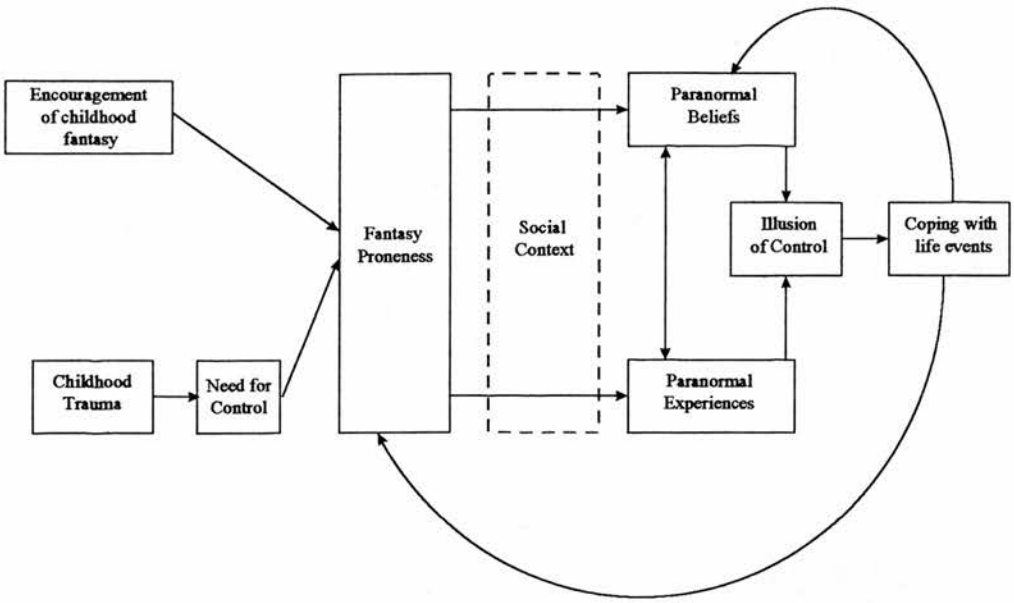

Figure 1. Irwin's (1992) model of the origins and functions of paranormal belief.

From figure 1 we can see that Irwin's model clearly addresses the developmental history of the paranormal believer, and some aspects of psychological functioning that may be deemed to be a consequence of paranormal belief. Clearly there is a central mediating role for fantasy in the development of Paranormal Experiences and Beliefs as far as Irwin is concerned, though Irwin sees this role itself as being influenced by a large range of social factors, his social context. Moreover, fantasy is seen to be equivalent with respect to its influence on both paranormal belief and experience. As for paranormal belief and experience themselves, Irwin sees them as correlating with one another, and thus seems shy of actually causally prioritising one or the other. In fact, no causal relationship between belief and experience is implied at all by Irwin's model. Nevertheless, Irwin's model does place paranormal belief and experience in to a model that shows some promise in the explanation of many key findings in research in to paranormal belief. For example, the cognitive deficits hypothesis implies an essentially attributional role for such deficits in the promotion of ersatz paranormal

experiences which then form the basis of paranormal beliefs, which themselves form a yet firmer basis for the promotion of yet more paranormal experiences. The source of paranormal experiences is attributional in that nothing in the experience in itself needs to appear paranormal.

For example, deficits in one's appreciation of chance events might lead one to suppose that the receipt of two identical birthday cards is a minor miracle, demanding some synchronistic or telepathic collusion between two independent senders. In fact such an event may be much more likely than is supposed. What is interesting here is that this kind of experience is constructed out of a set of independently mundane facts of the world. The unusual nature of the two card event is derived purely as an epiphenomenon of certain mundane events being coupled with a biased interpretation of them. However, commonly when someone claims to have seen a ghost this is precisely not the way such an experience could have occurred. Ghosts and apparitions, if seen clearly (as most seem to be), are seen as a distinct objects in one's field of view, that simply could not have been there (Tyrell, 1949; Wilson, 1995). Ghosts are thus truly experiential in a direct way; a way in which the birthday card experience can not be.

So, to bring this rather elliptical discussion to its conclusion, what is interesting about Irwin's model is that, with its emphasis on fantasy as the central mediating variable in the development of paranormal belief and experience, it appeals directly to the ghost-esque view of paranormal experience; the view that implies an experiential source for belief in the paranormal (Hufford, 1982). To believe in the paranormal and base this belief on experience then, the believer must have had a paranormal experience of the ghost-like kind, and not merely the birthday kind. Indeed, an experience of the birthday kind may well be something entirely consequential, happening only after one had developed paranormal beliefs. Further still, the (early) life long developmental view that Irwin's model encourages has the potential to account for all the other main

theories of paranormal belief (cognitive deficits, marginality view, personality/psychopathology view). Trauma is clearly implicated in the promotion of psychopathology, and must also surely have effects on childhood educational achievement, and for that matter one's social status.

I think for these reasons then, that Irwin's model represents a major step in the development of a truly useful model of the development not just of paranormal belief, but of paranormal experience also. As a consequence, I devote the greater part of this thesis to an attempt to confirm and extend Irwin's model. It should now be clear that research in to paranormal belief and experience suffers from precisely those problems I outlined in the beginning of this introduction. As I have said, what is required if research in this area is to progress from a merely descriptive stage is a decent model of the development of paranormal belief and experience and a method for sensitively and rigorously testing this model. Irwin has provided the basis for this model, and in the next chapter I outline the method I shall use to try and achieve to provide it with a sensitive and rigorous test.

Chapter 2

COVARIANCE STRUCTURE MODELLING IN THE ANALYSIS OF PARAPSYCHOLOGICAL SURVEY DATA: AN INTRODUCTION

Covariance structure modelling (CSM) represents a relatively new, powerful statistical technique for the analysis of correlational data, which should be of special interest to parapsychologists conducting surveys of paranormal experience and belief. With the capability to combine factor analysis, regression analysis, and path modelling, this technique has gained prominence amongst psychologists, sociologists, and other social scientists concerned with drawing conclusions from complex multivariable datasets. In particular, the capacity for CSM to aid the development of rigorous, theory driven approaches to survey data is emphasised.

Parapsychologists, like other behavioural scientists, have often turned to surveys to examine the psycho-social factors underlying people's experience of, and attitude towards, paranormal phenomena as will be abundantly clear from the previous chapter. The general need to understand the nature and function of paranormal experiences and beliefs in everyday life, has meant that a considerable body of survey based research, conducted by both parapsychologists and sceptics, has grown up over the past three decades. Typically these studies have employed simple regression and correlation techniques in order to assess the relationships between the variables of concern (e.g. Irwin, 1985; Thalbourne, 1984; Glicksohn, 1990). However, in contrast to experimental techniques, a common problem with surveys in general, and correlational survey data in particular, is that such studies often seem unable to tell us anything of genuine interest, beyond the fact that variable x seems to be correlated with variables y_1, $y_2 \cdots y_n$, to a more or less significant degree. In short, the old problem of inferring causes from correlations raises its head.

Arguably as a result of the relatively unsophisticated way in which contemporary survey type research has been conducted, it has largely failed to enhance our understanding of the nature and functions of paranormal belief and experience. Although the number of these studies is quite high, the field seems to have developed no theoretical consistency. Indeed, all too often correlational studies suffer from the researcher's piecemeal and somewhat ambiguous interpretation, based upon visual inspection of the correlation matrix or at best some type of hierarchical multiple regression analysis (to his credit Irwin, 1992, has actually been quite successful using just this type of approach).

That our understanding of paranormal experiences and beliefs should falter as a result of this kind of 'intuitive causal modelling' approach should come as no surprise. The answer to our problems lies in finding a technique which combines the theory driven, causally oriented, rigour of the laboratory experiment, with the convenience, diversity, and flexibility of the survey study. Fortunately, such a technique already exists, although to date its presence has gone unnoticed amongst survey researchers in the field, and it is to a discussion of the utility of this technique that I now turn.

Covariance structure modelling in the behavioural sciences

Until the late sixties, regression analysis and factor analysis constituted the two main ways of treating correlational data, and the choice of either one was largely constrained by the goals of the investigation. If one wanted to predict some variable's outcome in terms of one or more others, then regression analysis was used. If one's aims were to reduce a complicated multi-variate dataset to a smaller set of underlying or 'latent' variables then factor analysis was used.

However, there arose amongst social and behavioural scientists a general need to combine both the descriptive and predictive power of the regression

analysis, with the summarising power of the factor analysis. In particular it was felt that there were many weaknesses of the former that could be solved by the latter, and vice-versa. Consequently, over the last twenty years a new approach to the analysis of correlations, known variously as Structural Equation Modelling, LISREL, or Covariance Structure Modelling, has revolutionised the treatment and interpretation of correlational data. Amongst other things, this approach (which from now on I shall refer to as CSM) constitutes an extension of path analysis (a technique for analysing a hypothesised causal path underlying correlational data, often using hierarchical multiple regression analysis). With CSM a factor analytic model (the measurement model) is used to assess which variables load on which factors, and a regression type model (the structural model) is used to assess the interrelations between factors, in order to test these interrelations against the hypothetical causal model input by the researcher. This approach thus promotes the rational, statistical testing of hypothesised causal relations between data from correlational studies. In detail the approach is highly mathematical, requiring a good general understanding of matrix algebra in order to really follow the mechanics of the technique. However, given my own lack of familiarity with all but the basics of matrix algebra, and the constraints of space, I shall employ an intuitive, example based approach to covariance structure modelling, showing concretely how CSM can significantly enhance the validity and utility of studies using correlational data. Those who wish to delve into the mathematics of CSM may find Long (1994), Bentler (1989), and Bollen (1989) useful. A particularly good introduction to CSM (using Bentler's EQS program) for non-mathematicians is Byrne (1994).

Covariance Structure Modelling in Practice

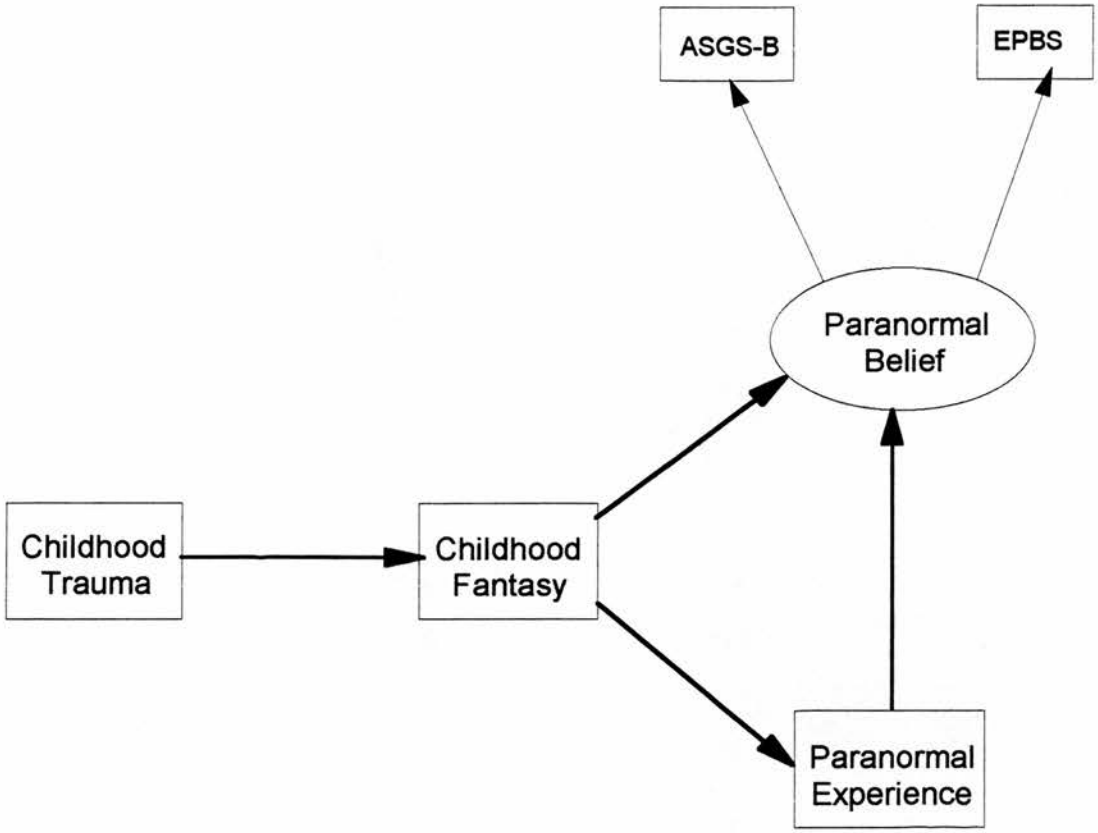

Figure 1. A path diagram of Irwin's (1992) childhood factors model.

When using CSM to test causal models[1] on some particular data set the problem arises as to what model to test. If our research is pioneering, like we all think it should be, then there may be no explicit models or theories to account for the findings in the field being investigated. In this case we just have to plough through the studies and the reviews and come up with a naive model of what's going on. This model may bear close relation to folk psychological notions (i.e. we may be using common sense to guide our theory). This is fine, as the important thing is to have something to test (even if its wrong).

[1] No amount of statistical sophistication will glean causation from correlation. The imputation of causality within a causal model stems from the background research and theory underlying the model, and does not reside within the technique itself.

However, usually when we come to attempt CSM we do so because some basic model has been postulated to account for the phenomena in which we are interested. Given that a basic model exists, the first thing to do to turn it into something suitable for use with CSM is to draw it as a graphical model (i.e. set up a path model specifying the causal relationships between the variables modelled). In figure 1. we can see a graphical model, produced by Irwin (1992), which attempts to outline the early childhood causes of paranormal belief and experience. How can we use CSM to test Irwin's model? In what follows I shall guide the reader through the process of specifying Irwin's model in the language of CSM used by EQS.

According to Irwin's path diagram childhood trauma causes childhood fantasy (as a means of escape from a harsh reality), and childhood fantasy leads to both paranormal experience and paranormal belief. In addition, paranormal experience causes paranormal belief. A basic premise in EQS is that any variable said to be caused by another variable can be expressed as a linear function of the causative antecedent variables. Thus, if we say that trauma causes fantasy, we are saying that fantasy proneness scores can be expressed as a linear function of trauma scores (plus some error) thus;

fantasy=trauma+random error.

but as EQS refers to variables as V's and error terms as E's then we can reexpress this equation in EQS speak as;

*V2=*V1+1*E2.*

where V2 is fantasy and V1 is trauma, E2 is error associated with the prediction of fantasy.

Readers familiar with simple linear regression will see that this equation is broadly similar. A full model is 'specified' by providing a series of equations of varying complexity, like the one above, for all of the features of the model that demand specification. The trick in specifying a model is to know which parts of

the model should be freely estimated and which parts should be fixed (usually at zero, and in that case this is done by not specifying them at all). Relationships in a model which are to be estimated are prefixed by an asterisk. Thus the relationship between V1 and V2 above is to be estimated because the equation contains the expression *V1 rather than 1V1 (fixed at 1) or 0.5V1 (fixed at 0.5). In contrast, the relation between V2 and E2 is fixed at 1, because it makes sense to suppose that V2 is responsible for all of the error associated with its own prediction[2]. Once Irwin's model has been fully specified it might look something like this;

```
/EQUATIONS
V6 =  + *V9 + *V13 + E6;
V7 =  + 1F1  + E7;
V8 =  + *F1  + E8;
V9 =  + *V13 + E9;
F1 =  + *V6 + D1;
/VARIANCES
V13 = *;
E6 = *;
E7 = *;
E8 = *;
E9 = *;
D1 = *;
/COVARIANCES
```

[2] This will not always be true. Sometimes errors between variables are correlated, and Bentler (1989) discusses some interesting cases where error terms were themselves modelled as causally efficacious on other normal variables.

There are some other features of the above that need to be noted. Thus, we can see the statement F1 = + *V6 + D1. This refers to a causal effect of a measured variable upon a latent variable (one that is made up of a number of measured variables and is thus a factor - F). Factors in covariance structure models have error terms referred to as 'disturbances' (D). We can also see that V7 (EPBS) has a loading fixed at 1.0 on its factor (F1 - paranormal belief). This is done to 'scale' that factor[3], and it also aids in identifying the model mathematically. A model is said to be identified when only a single set of parameter values can be obtained from the observed covariances and variances of the original variables. Identification is related to the notion that a simple equation in two unknowns can not be solved with only a single piece of information (data), e.g. A*B=1 cannot be solved because their are two free parameters A and B but only one data point. If we add another equation involving B though, say B=3, then the parameters become just identified.

This model incorporates a very straightforward conception of the covariance structure model, known as the Bentler-Weeks representation. Thus, according to the Bentler-Weeks representation a model variable is said to be dependent when it is depicted as a linear function of any other variable in the model (i.e. is placed on the left hand side of a linear equation in the equations section of the EQS model file. Conversely, a variable is said to be independent when it is not specified as a linear function of any other variable (i.e. it exists only on the right hand side of a linear equation in the equation section of EQS). Thus, the variances and covariances of a model are also, in the Bentler-Weeks representation, independent variables. This conception of the CSM is straightforward in that it describes models in terms that any one familiar with regression analysis would understand, as opposed to the jargon ridden matrix algebra conceptualisation favoured by proponents of LISREL. Using a model like this EQS is able to specify a structural matrix of coefficients that represent

[3] Equivalently one could fix the variance of the factor to 1.0, a more appropriate technique when confirmatory factor analysis is used and the precise value of factor loadings is required.

the theoretically relevant relationships in the model. Without getting too heavily in to the matrix algebra, what EQS does is attempt to reproduce the observed correlation matrix O using just these coefficients and a technique for manipulating their values. The immediate end product of an EQS analysis is an estimated correlation matrix S which when subtracted from the actual correlation matrix O gives an intuitive measure of the degree to which one's model reproduces the observed data[4]. Most of the other fit statistics that EQS produces are based upon calculations drawn from this residual correlation matrix.

A variety of specialised programs exist that can perform CSM. Generally the technique is too complex to be performed on general statistical packages. The

$$\begin{pmatrix} 1.0 & 0.5 & 0.3 \\ 0.5 & 1.0 & 0.4 \\ 0.3 & 0.4 & 1.0 \end{pmatrix} \quad \begin{pmatrix} 1.0 & 0.45 & 0.1 \\ 0.45 & 1.0 & 0.39 \\ 0.1 & 0.39 & 1.0 \end{pmatrix} \quad \begin{pmatrix} 0.0 & 0.05 & 0.2 \\ 0.05 & 0.0 & 0.01 \\ 0.2 & 0.01 & 0.0 \end{pmatrix}$$

$$O \quad - \quad S \quad = \quad R$$

Figure 2. Example of subtraction of estimated correlation matrix S from observed matrix O to give residual matrix R.

best known CSM packages are LISREL (Joreskog and Sorbom, 1981), MILS (Schoenberg, 1982) and EQS (Bentler, 1989). Of these LISREL is the most famous, MILS the cheapest, and EQS the most user friendly. EQS combines factor analysis, multiple regression, and path analysis techniques, providing a powerful system to test and confirm various causal or structural models. Traditionally covariance structure modelling involved the arduous specification of the model using a tricky programming language to specify the various measurement and structural (i.e. model) equations that form the model to be tested. Fortunately, the latest version of EQS has an 'easy build' option which brings causal modelling within reach of those not maths minded (although one

[4] Though the actual measure of fit used to determine the success of any modelling attempt may vary

can still specify the model directly). Our model is then converted into the language of mathematics that EQS understands, and a hefty dose of matrix algebra later we have some output.

Whilst the maths involved is too complicated for a general introduction, the basic principle behind CSM is easy to grasp. A researcher gathers together data from a variety of measured variables, and puts these into a matrix of intercorrelations. The researcher is concerned to explain the observed relationships between the variables in terms of some theory. This theory, when specified appropriately, can be used to construct a second matrix of estimated intercorrelations. The aim of CSM then is to find a system of linear structural equations, in the form of a model, which are able to reproduce as much of our observed correlational data as possible. An estimated correlation matrix which accounts for almost all of our observed correlations is said to be a good 'fit'. One which leaves much of the variability in our data unexplained is said to be a poor 'fit'.

Consequently, the most important output from a test of any causal model is a list of goodness of fit statistics, indicating how well the hypothesised model accounts for the observed correlational structure of the data. When assessing the goodness of fit of a causal model there are a number of statistics which we have at our disposal. For the first of these, the chi square goodness of fit statistic, we require that the chi square be *non-significant* for our model to be considered a good fit. Next we have the residual statistics, of which the off diagonal average standardised residuals are the most important. These may be treated like residuals in regression analysis, indicating unexplained variance, and are particularly useful in comparing the differences in fit between two or more alternative models. Aikaike's Information Criterion (AIC) (Aikaike, 1987) chi-square provides an estimate of goodness of fit that takes into account the parsimony of the underlying model (in terms of the number of parameters that

---

according to the fit optimisation technique used (e.g. ordinary least squares, maximum likelihood).

must be estimated to arrive at a particular solution). Thus the smaller the value the better. Lastly we have the fit indices. Of these, Bentler and Bonett's non-normed fit index (NNFI) (Bentler and Bonett, 1980) provides a particularly good estimate of model fit with small sample sizes (appropriate for our example sample size). To even be considered as a reasonable fit, any fit index should show values of greater than 0.9.

In addition to these fit indices, a useful feature of causal modelling is the suggestion of model modifications, whereby EQS suggests changes to the model in order to provide a better fit. Used cautiously, this feature is helpful in achieving models which fit the data well. Thus, EQS provides the Wald test for dropping parameters which can aid the user in removing non-significant or non-important parameters from a model, whilst the Lagrange multiplier test suggests individual or multivariately significant parameters which if added in to a model might substantially improve model fit. For cautionary discussion of these modification indices see MacCallum, Roznowski, and Necowitz (1992).

We should note at this juncture that covariance structure modelling is not viewed rosily by all. Biddle and Marlin (1987) present a number of reasons why CSM should be approached warily. Firstly, they raise the old cause-correlation problem, arguing that all too often researchers speak as though one wave of the CSM wand over their correlational data will magically reveal causal relations. This argument holds a certain degree of truth, although it is often used too liberally to entirely discredit the CSM approach. Thus while it is certainly the case that mere correlations do not imply causations, it is also quite correct to note that causations always imply correlations. The question then becomes one of establishing, by means which lie outside of any particular analytical technique, the a priori plausibility of the causal statements derived from a model. As always, data by themselves say little outside of the context of a background theory. One important point made by Biddle and Marlin is that the weakest form of 'causal modelling' is that which is based on cross-sectional, single wave field studies. Given time constraints that is the only kind of study

we will find in this thesis. Biddle and Marlin also criticise the misuse of the term confirmatory when used to describe the aims and goals of CSM. Specifically, they argue that researchers often confuse failure to reject a causal model with 'proof' of that model's veracity in a way which leads them to suggest that the model has been confirmed. Their arguments tell most strongly against claims for the absolute confirmation or 'proof' of a model, but are very much weakened in the case where relative confirmation of model fit is sought against one or many competing models. The testing of a preferred model against competing alternatives is regarded as good practice by MacCallum, Wegener, Uchino and Fabrigar (1993).

Biddle and Marlin also point out that the use of path diagrams with causal models is often confusing, and that as yet no conventions regarding the construction of these diagrams is forthcoming (though there is more consistency now than there was at the inception of the path diagram).

CSM with latent variables (variables derived from the factoring of manifest variables which we directly measure) also concern Biddle and Marlin, in the sense that path coefficients based on latent variables often have larger coefficients than equivalent models based on measured variables alone. Biddle and Marlin agree with Bookstein (1984) that this is like '...making "something out of nothing,"'. However they note that their position on latent variable modelling 'enhancements' depends upon the stance one takes with respect to measurement error in regression coefficients. My own understanding is distinctly less sceptical than Biddle and Marlin's. I was brought up to understand that, providing one's measured variables were reasonably reliable and valid, factoring them in to a latent variable would remove the specific systematic and random error variance of each, thus giving a cleaner variable with less diluted coefficients. Nevertheless, one must be aware of the differences that arise in estimation of path coefficients between ordinary least squares CSM and other techniques. Lastly, Biddle and Marlin point out the expense and time cost of CSM techniques. However, in the ten years since

their paper was published the advent of Pentium PC's has made CSM something within the computational reach of almost everyone (the longest analysis to be reported in this thesis using CSM was around 30 seconds, most took no more than ten seconds!).

Despite Biddle and Marlin's caveats about the use of CSM, the advantages for theory development and analysis of survey data using this technique offer a considerable improvement over previous analytic strategies. Unlike previous correlational work in the area of paranormal belief, which has tended to be weakly theoretical and hopelessly fragmented, CSM offers the possibility of coherent, rigorous, single stage, multivariate tests of precisely specified hypotheses concerning paranormal belief and experience. We will never be able to make the jump from correlational to causal using such techniques, but we can at least hope to assess the causal relevance of correlational data in an organised manner.

Chapter 3

HOW MANY FACTORS OF PARANORMAL BELIEF ARE THERE?: A CRITIQUE OF THE PBS[1]

In this chapter I take something of a detour from the main empirical path of the thesis. This chapter started simply with the idea of surveying the range of measures of paranormal belief, and ended up with an academic debate with the main author of one of the most widely used measures of that construct, Jerome Tobacyk. By the time 'battle' had been joined it became clear that I would have to devote some additional empirical effort to resolving our debate, and so two large scale confirmatory analyses were conducted which largely supported my critique. They are also included as an integral part of the thesis.

Psychological research on belief in the paranormal has important consequences for our understanding of these commonly held views and the people who hold them. But research conducted on belief in the paranormal is only as good as the devices used to measure paranormal belief. There are numerous questionnaires available to assess paranormal belief (Jones, Russell, and Nickell, 1977; Randall and Desrosiers, 1980; Scheidt, 1973; Thalbourne and Delin, 1993; Tobacyk, 1988). However, the most popularly used measure of paranormal belief is Tobacyk and Milford's (1983, 1988) Paranormal Belief Scale (PBS). Perhaps the PBS' greatest contribution to the area of paranormal belief measurement is its emphasis on the multidimensional nature of paranormal belief. For early researchers, belief in the paranormal was a unitary phenomenon—a generalised trait much like intelligence—and unlike personality (e.g. Randall and Desrosiers, 1980, found a single main factor of 'supernaturalism' accounted for 70% of the variance in their factor analysis).

---

[1] This chapter appeared as a published paper in the Journal of Parapsychology, followed by a counter reply by Tobacyk (1995a;1995b).

But despite early enthusiasm for a unidimensional construct of belief in the paranormal, later factor analyses of paranormal belief have tended to support the view that belief in paranormal phenomena is multi-faceted (Sobal and Emmons, 1982; Grimmer and White, 1990; Clarke, 1991; Thalbourne and Delin, 1993).

Indeed, Tobacyk and Milford's (1983, 1988) PBS has seven dimensions of paranormal belief. These subscales are (in descending order of percentage of variance accounted for); Traditional Religious Belief, Psi Beliefs (mostly Psychokinesis questions), Witchcraft, Superstition, Spiritualism, Extraordinary Life Forms, and Precognition. Since the development of the original scale a number of changes led to a revised scale (Tobacyk, 1988). The revised PBS incorporates a 26 item, seven point scale (as opposed to 25 item five point scale) and has changes to the Precognition scale, Witchcraft scale, and minor changes to the Extraordinary Life Forms scale.

The arguments expressed in this chapter stem from the suspicion that seven dimensions of paranormal belief is simply too many. Whilst several researchers (e.g. Irwin, 1993) have raised specific concerns about the PBS, no one has as yet actually examined in depth its construction and content, to see if the claims for seven dimensions are accurate. Consequently, in the remainder of this chapter I present an in depth critique of the PBS, and its revised variant. In the first section I attempt to determine precisely how the factor analysis of the PBS could have arrived at seven factors. Factor analysis is a complicated statistical technique, and in reconstructing the original factor solution of the PBS the technical details presented may be skipped by those who are of a less technical persuasion. These readers are advised to move on to the last (summary) paragraph of the next section, and then to read on from 'The Revised PBS'.

The development and factor analysis of the PBS

In constructing their paranormal belief scale Tobacyk and Milford (1983) gave sixty-one items that in their judgement "...would sample as wide a range of paranormal beliefs as possible." (pp. 1030) to 391 students at Louisiana Tech University. Items were scored on a five point scale from 1 (strongly disagree) to 5 (strongly agree). These items were intercorrelated and factor analysed by the principal axis method. Promax oblique rotation was performed to assess for dependencies between factors but "revealed little evidence of dependence between factors, which allowed for the use of orthogonal rotations." The orthogonal transformation solution gave 13 factors, accounting for 60.7% of total variance.

Tobacyk and Milford (1983) do not state whether any statistically based means of factor selection were used to arrive at the original 13 factor solution, except that they go on to select only 7 of the 13 factors on the basis that these factors '...appeared meaningfully interpretable as common factors.' However, Dr. Tobacyk kindly informed me that, 'I used several criteria for the selection of factors concurrently: eigenvalue > 1.0; scree plot of eigenvalues, proportion of total variance accounted for, conceptual meaningfulness of the factor dimensions' (personal communication, 1994).

It would be worthwhile to note at this juncture that Tobacyk's set of four criteria for selecting factors, whilst reasonable in an early perusal of the analysis, fall short of adequacy in the construction of the final scale. The trouble is that these four criteria, taken together, are likely to produce far too many factors. Also, whilst selection of factors that "...appeared meaningfully interpretable" is a valid procedure (providing that one is clear about what would constitute a meaningful factor *a priori*), in Tobacyk and Milford's case use of this criterion actually contradicts their explicit aim to let the factor analysis describe the structure of paranormal beliefs, rather than impose any *a priori* dimensions upon the data (a la Randall and Desrosiers, 1980). Indeed, it is hard to see how one can select factors that appear meaningfully interpretable without having

some *a priori* conception of what should constitute a factor of paranormal belief. With at least four criteria for selecting factors from their analyses, Tobacyk and Milford's factor selection procedure may have assumed the characteristics of a projective test.

We can find out just how many factors there would most likely be had only one, statistically reasoned, factor selection procedure been used. To start with, we can reject the simple use of the Kaiser-1 criterion on the grounds that it is too liberal (see Horn, 1965; & Zwick and Velicer, 1986), giving too many factors, and indeed, in Tobacyk and Milford's analysis, it would have accounted for 13 factors alone[2]. Alternatively, looking at the proportion of variance accounted for by each variable is a redundant procedure, as an eigenvalue is simply an indirect measure of proportion of variance accounted for, so any procedure that analyses eigenvalues to select factors is 'looking' at proportion of variance accounted for. Also, we cannot use the most accurate selection procedures, such as Minimum Average Partial Technique (MAP) or Parallel Analysis (PA), as these techniques require access to original data, which Tobacyk no longer has (Tobacyk, personal communication, 1994).

We are thus left with the use of a scree slope analysis which involves plotting a *visually judged* line of best fit through the smallest eigenvalues, and taking only those which come above the line In order to do the analysis, we need to reconstruct the 13 factors of the original factor solution. We know the eigenvalues of the first seven factors from data presented in Tobacyk and Milford (1983), and by making certain assumptions we can piece together the likely distribution of the original 13 factors. The assumptions that permit us to do this are as follows:

1) *Factor seven (precognition) is the seventh largest factor*- we are given no reason to believe that Tobacyk and Milford did not select the seven largest

[2] We can reasonably assume that the first 13 factors stemmed from the default application of Kaiser-1 by Tobacyk and Milford's SAS statistical package.

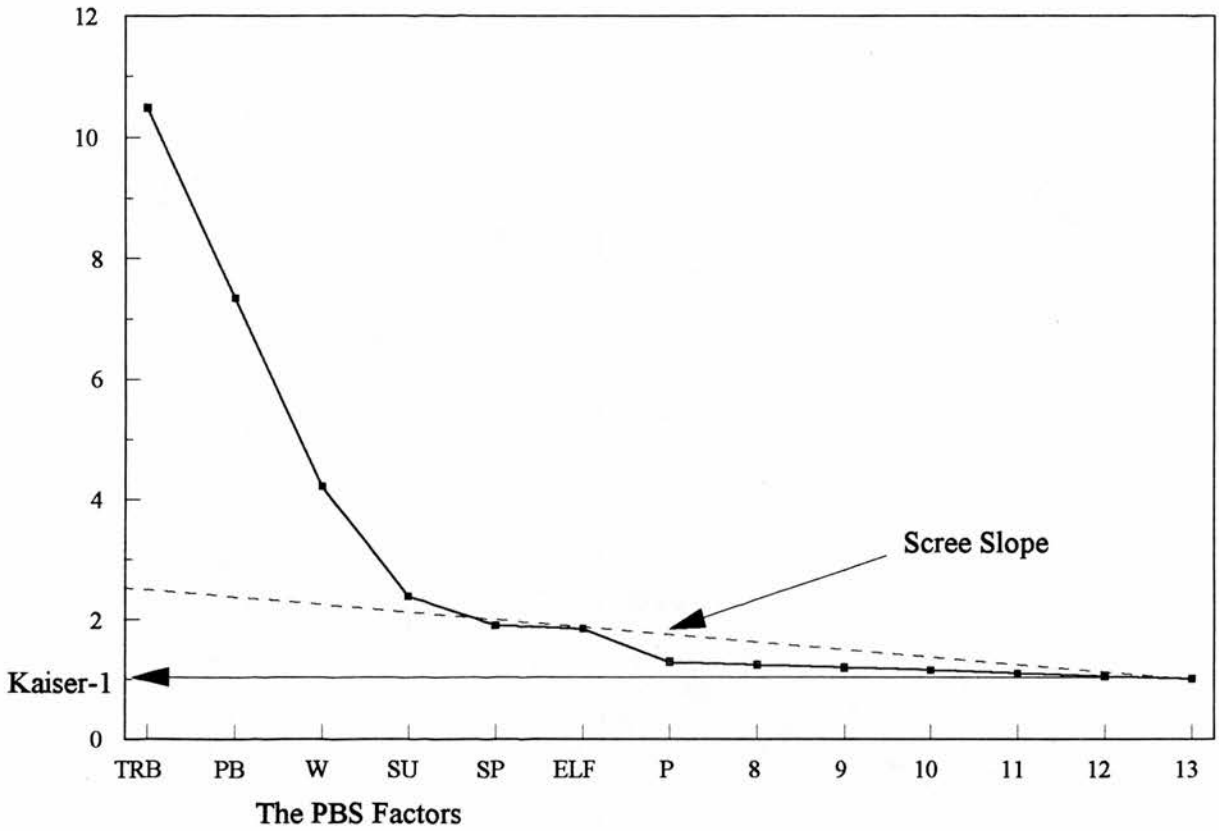

Figure 1. A plot of the thirteen factors of Tobacyk and Milford's (1983) Paranormal Belief Scale.

factors from their 13 factor solution, and it seems unlikely that they did otherwise.

2) *Factor thirteen had an eigenvalue of exactly one-* if this was the smallest factor to make the Kaiser-1 'cut' then it is most likely that it was on or very near the criterion division of one.

3) *Factors 8 to 13 were roughly linearly distributed (i.e. triangular)-* this assumption is not unreasonable. It is the weakest assumption I make however.

We know that Tobacyk and Milford's original thirteen factor solution accounted for 60.7% of the total variance in their analysis. And, by summing the percentage variance accounted for by each of their seven selected factors,

we know that these seven factors accounted for 48.2% of the total variance. This means that, of the thirteen original factors, the six 'unknown' factors that did not make the cut accounted for 12.5% of the total variance. By invoking assumption 1 we can deduce that Factor eight had an eigenvalue no more than that of Factor seven (Precognition). In this case that is 1.281 (2.1/100*61). So Factor eight must be less than 1.281. From assumption 2 we can suppose that the smallest of the thirteen factors selected had an eigenvalue very close to one. However, I shall suppose that the eigenvalue for Factor thirteen was in fact one (in any case these calculations are unlikely to be far out). Assumption 3 may be invoked to determine the distribution of the unknown 'six' factors. To do this we simply suppose the distribution of these factors to be linear and work out the gradient of the slope for these factors using basic mathematics. In this the 'distance' the slope must travel is six factors, and the height the slope must reach is from an eigenvalue of 1 to an eigenvalue of 1.281 (or 0.281 units). The gradient for this slope is given by 0.281/6 = 0.0468166. This value is then the average distance between factors seven to thirteen. So using those first three assumptions we can now get some idea of the distribution of Factors eight to thirteen. The results of this analysis, expressed in eigenvalues, along with the known eigenvalues of the seven factors of Tobacyk and Milford's PBS, are given in figure 1 (N.B. Factors eight to thirteen are inferred).

Now that we have a rough plot of the eigenvalues in Tobacyk and Milford's PBS we may try to use the more accurate and discerning scree slope criterion to select out the factors in their analysis. Using just the scree slope analysis[3] they would have selected Traditional Religious Belief, Psi Belief, Witchcraft and possibly Superstition (depending on where exactly one supposes the line to be a best fit). So on these most basic analyses there are only four factors at most.

[3] Some earlier commentators on this chapter have expressed doubts about the accuracy of the scree test over that of the Kaiser-1 criterion. However, Zwick and Velicer (1986) found it to be, on average, 35% more accurate than the Kaiser-1 criterion in selection of predetermined factors (57% vs. 22% respectively). Earlier reports also attest to a high degree of accuracy (Cliff, 1970; Cattell and Jaspers, 1967; & Zwick and Velicer, 1982) and generally good interrater reliability among expert and naive judges (Cattell and Vogelman, 1977; Zwick and Velicer, 1982).

The other factors might be meaningfully interpretable, and have eigenvalues above one, but they do not contribute much understanding to the concept of paranormal belief over the first three or four factors, nor for that matter do they express much more summarising value than any one of the original 61 items. Indeed, there are conceptual reasons to suppose that the precognition factor should not be a separate factor at all, but a part of psi belief as it surely is [see later]. I shall now move on to the second major area of problems with the development of the PBS scale, where I discuss in particular concerns about the smaller factors.

The revised PBS

In 1988 Tobacyk produced an as yet unpublished paper outlining changes made to the original PBS, and declared on this basis a new revised PBS. The changes made are summarised as; (1) original 25 items made up to 26, (2) new seven point scale, (3) replacement of original 3 item Precognition scale with a four item scale of all new questions, (4) replacement of two of four Witchcraft items, and (5) replacement of one of three Extraordinary Life Form items.

It is argued that the changes result in greater reliability and validity, less restriction of range, and greater cross-cultural validity in Western cultures. The inclusion of a seven point rating scale seems fairly straightforward and makes good sense. In what follows I shall address problems with the various subscales of the PBS, before going on to consider whether Tobacyk and Milford's early evidence for construct validity is acceptable. Some of these problems are minor, and can be easily remedied. Some however, threaten the very existence of the PBS. First I begin with the second most important factor of the PBS, the Psi subscale.

The Psi subscale

Parapsychologists define 'psi' as an unexplained human extra-sensorimotor capability comprising telepathy, clairvoyance and precognition, for ESP; and

psychokinesis (mind over matter) for PK. Thus as a minimum criterion for that name, any psi factor should meet the requirement that its items broadly and comprehensively deal with belief in telepathy, clairvoyance, precognition, and psychokinesis. Yet the PBS Psi belief factor has only four questions; of which three address PK belief, and one refers to telepathy alone. Clairvoyance is not addressed, and precognition appears on its own as the smallest factor. Plainly, the so called Psi belief scale is poorly named - it is, for all intents and purposes, a PK belief scale. The second most important factor of the PBS is simply too shallow to warrant its title, or its importance. Lest it be renamed the 'PK belief factor' a new Psi belief scale is needed.

Precognition subscale

Precognition is a particular kind of psi phenomenon that involves psychic perception of a future event, and along with clairvoyance and telepathy is characterised and commonly understood to be a form of extra-sensory perception. Thus one might expect people who respond affirmatively to questions about belief in telepathy to also respond affirmatively to questions about belief in precognition, and vice versa. In short, responses to questions on precognition belief should correlate positively with those on other forms of extra-sensory perception. Thus, there are sensible *a priori* reasons to suppose that belief in telepathy, clairvoyance and precognition should be expressible in terms of a common factor of ESP beliefs (and to a lesser extent for PK beliefs). That the PBS showed precognition to be a separate factor from psi beliefs is perhaps the single most puzzling outcome of Tobacyk and Milford's factor analytic assessment.

It is notable that for the revised PBS (1988) Tobacyk totally replaces the three questions on precognition with four new ones (hence accounting for the increase of PBS items from 25 to 26). The reasons for doing this are outlined by Tobacyk as; 1) the original Precognition subscale has the lowest test-retest

reliability of the original seven subscales (*rtt* = .60 over four weeks)[4]. 2) The original questions were ambiguous and could have referred to non-paranormal means of predicting the future. Tobacyk himself gives good reasons to believe that all three original questions suffer from ambiguities of this sort. Thus in the original PBS the ambiguity of the questions would likely have meant that a large number of respondents were not viewing these questions as being about "psychic prediction of the future" but merely about the likelihood that people could predict the future in general (say perhaps with meteorological techniques). And obviously these responses are not likely to covary meaningfully with the psi scale items. In fact they are not likely to covary meaningfully with any of the sub-scales in the PBS. These questions are just going to sit as a separate factor on their own. Which is precisely what they do.

It is clear then from Tobacyk's (1988) own paper that the Precognition sub-scale was far from satisfactory: both in its item content and reliability. The key issue was that the items that went in to the original factor analysis related to the concept of paranormal belief only ambiguously. Tobacyk's (1988) solution to this problem is to replace these ambiguous questions with questions that are obviously related to precognition. These new questions were taken from a study (Tobacyk, Nagot and Mitchell, 1989) that had 349 college students rate 60 beliefs about the prediction of future events. A factor analysis of these sixty items revealed seven independent dimensions of belief about the prediction of future events. According to Tobacyk (1988), "The two clearest marker items were selected from each of the two largest paranormal precognition dimensions to form a new Precognition subscale. These dimensions and items were: Paranormal Divinatory Systems (Astrology is a way to accurately predict the future. The horoscope accurately tells a person's future.) and Psychically Gifted Persons (Some psychics can accurately predict the future. Some people have an unexplained ability to predict the future.)." (pp. 6).

[4] Kline (1986) notes that for a minimum acceptable level of reliability one should try to get at least a retest reliability of .7.

Thus from this study the four questions above were used to form the new Precognition subscale. The test-retest reliability for this scale is a reasonable .81 over four weeks. However, there are numerous errors and flaws with the approach to the new Precognition subscale that must render it invalid as a separate scale altogether. I shall now outline these problems with the Precognition subscale in the revised PBS.

1. *Original evidence for a factor of precognition was flawed* - As Tobacyk (1988) agrees the original items in the PBS Precognition subscale were ambiguous. Therefore the possibility that the Precognition subscale does not actually measure belief in paranormal precognitive powers must be taken seriously. If evidence for a separate dimension of belief in precognition depends upon such ambiguities then this raises serious questions about the validity of the precognition dimension. A separate factor of precognition can not be upheld if the items that contributed best to that factor do not actually relate in an obvious way to precognition. Evidence for a separate factor of precognition must be derived anew. So an error is made in supposing that whilst the items composing the original Precognition subscale were in need of change the actual status of the factor of precognition needed no such scrutiny. This supposes that factors are more than the sum of their parts. In fact the existence of factors depends crucially upon the quality of their components- the subscale items come first, the factor is always a latent variable that is to be inferred.
2. *The factorial independence of the new precognition items is not established* - This criticism stems from what has previously been argued. One should really attempt to determine anew that precognition belief is a separate aspect of belief in the paranormal. Only then can the claim be made that a separate factor of belief in precognition exists, and only then can the new precognition items be placed into such a factor. Instead, in choosing to develop new PBS Precognition scale items only with

reference to items that have 'prediction of future events' content Tobacyk and his colleagues lose all possibility of establishing that a factor of belief in precognition does exist, apart from other kinds of belief in paranormal phenomena. And this is precisely what is needed to justify the inclusion of a separate factor for precognition in the first place.

3. *The Tobacyk, Nagot and Mitchell (1989) factor analysis contradicts the notion of a single factor of precognition* - The idea of having a single factor of precognition in the original PBS was surely that those items that were in the Precognition subscale correlated highly with one another, but not with anything much outside that subscale. Tobacyk and Milford's (1983) justification and use of orthogonal rotations implies this much. Basically the items chosen for the original Precognition subscale represented different ways of asking the question- Do you believe people can predict the future paranormally? This is reasonable if there is good reason to suppose that a single factor exists in the factor analysis. In their development of the Belief in Prediction of Future Events Scale (BPFE) Tobacyk et al. (1989) found seven independent factors of belief in the prediction of future events. These factors underwent orthogonal rotation, producing factors that shared the least amount of variance. Subsequently, Tobacyk et al. took the strongest two marker questions from the clearest two factors of the BPFE to form their new Precognition subscale. However, in doing this they destroy wholeheartedly the notion of a single factor of precognition. This is because the four new questions from the BPFE are taken from two independent factors (Paranormal Divinatory Systems and Psychically Gifted Persons). In fact two of the questions in the new Precognition subscale refer to belief in astrology and not precognition. Only one of the other two questions actually refers in a clear and definite way to precognition. In short the new 'factor' of precognition, despite its improved reliability, measures two different factors

altogether, only one of which assesses belief in precognition (Psychically Gifted Persons). Changes made to the Precognition subscale may have improved the scale's reliability, but only at the expense of a near total lack of validity.

The Precognition subscale: Conclusions

The conclusion that one must draw from an understanding of this section is that there is no good evidence for a single factor of belief in precognitive powers. Had the construction of the original PBS precognition items been conducted differently then it is most likely that no separate factor of precognition would have been found. But there may be another reason why the precognition subscale did not form a part of a wider psi scale. As noted previously, the Psi Belief subscale contained three questions on psychokinesis, and only one on ESP. Whilst precognition belief correlates highly with ESP belief in general, it has persistently correlated less well with PK belief in previous analyses, for reasons that should be intuitive. If items relating to Psi Belief had been truly comprehensive, the precognition items would most likely have loaded onto the factor of Psi Belief- as they have in almost every other factor analytic assessment of paranormal belief to date (see Sobal and Emmons, 1982; Clarke, 1991; Grimmer and White, 1990; Thalbourne, 1981 and most recently Thalbourne and Delin, 1993; but see also Thalbourne and Haraldsson, 1980).

The Witchcraft subscale

The Witchcraft subscale of the revised PBS consists of four statements obviously relating to witchcraft and black magic. The original Witchcraft subscale had been changed because two of the original items (Voodoo is a real method to use paranormal powers. There are actual cases of voodoo death.) were often not familiar to foreign respondents (specifically Finish, German and Polish). These items were replaced by the statements 1) There are actual cases

of witchcraft. 2) Through the use of formulas and incantations, it is possible to cast spells on persons.

Of these two replacement items the first is basically fine, if a little equivocal (one could support the assertion that there are actual cases of witchcraft and still suppose that magic doesn't really exist). The second statement is excellent and taps into the factor of Witchcraft without actually mentioning it explicitly, which is always a good way of constructing one's items. However, of the remaining two statements in the revised Witchcraft subscale (Black magic really exists. Witches do exist.) it seems clear that the statement "Witches do exist." is obviously true whether or not they really have magical powers and it is likely to be endorsed as such by sceptic and Wicca worshipper alike. Perhaps this statement could be changed for something else like "Witches possess genuine magical powers". For the most part however there are no particular problems of major significance with the Witchcraft subscale.

The Extraordinary Life Forms subscale

This subscale consists of the questions; (1) The abominable snowman of Tibet exists, (2) The Loch Ness monster of Scotland exists, (3) There is life on other planets. The items in this scale are simply of arguable paranormality. The mystery surrounding the existence (or not) of the Yeti and Nessie stems more from their apparent elusiveness, than it does from any intrinsically inexplicable process of nature. If these creatures exist then they do so as a matter of natural historical fact and are as unlikely to herald a new principle of science as they are likely to warrant discarding an old one. However, the third statement in the Extraordinary Life Forms factor is probably quite useless for another reason altogether. Anyone—even a Royal Society astronomer—could agree with the statement that there is life on other planets to some degree. The real question that taps into the Extraordinary Life Form dimension should not have been whether there was life on other planets but whether life on other planets was regularly making trips to our planet! I suspect that very few astronomers would agree with that contention. So at the very least the factor of Extraordinary Life

Forms, which accounts for the second least percentage of variance in Tobacyk and Milford's original factor analysis, rests upon shaky foundations, theoretically and methodologically.

The subscales of the PBS

If one wishes to develop a psychometric measure of anything of import it is usually the case that one needs a large number of items to sample the particular latent variable of interest, be it belief in the paranormal or extroversion. One needs an even larger number of items to begin to narrow down the best items to go in to the refined measure. A scale which needs only 26 items to sample no less than seven independent latent variables is probably not valid[5]. Indeed, we have seen that the Psi belief subscale suffers from a lack of comprehensive coverage of the core psi phenomena, showing instead a prevalence of items relating solely to psychokinesis. If four or five items can get to grips with the complex attitudes and feelings underlying belief in the paranormal then either belief in the paranormal is trivially simple or our approach to item selection is wrong.

Problems of subscale independence (orthogonality)

In this section I draw on evidence from the old and new versions of the PBS. In the original paper on the old PBS Tobacyk and Milford presented a matrix of intercorrelations between the subscales of the PBS that shed light upon the question of whether the subscales are truly uncorrelated. Clearly from examining this matrix there are some correlations large enough to warrant concern over subscale orthogonality. For example, Psi belief correlates .49 with Spiritualism, .40 with precognition (where orthogonality has already been questioned), and .34 with Witchcraft. Tobacyk and Milford (1983) dismiss these larger correlations, not by dealing with them directly, but by noting that "Because these correlations are based on a 424-subject sample, statistical

[5] This is a point well made to me in correspondence from Paul Kline, Professor of Psychometrics at the University of Exeter, England.

significance alone is not an appropriate procedure to determine relatedness between subscales....The amount of shared variance (r^2) between scales appears to be a more appropriate index of relatedness." (p. 1034).

Certainly even the most minor correlations are likely to be significant with such a large sample size and r^2 is a useful estimate of the import of the relation between two variables, but one needs to be very wary indeed in throwing out evidence for relatedness (obliqueness) of subscales when the subscale correlations approach .40 and above. This wariness is especially advised given that subscale intercorrelations can only be as high as subscale reliabilities will allow. Indeed, it is a general psychometric principle that a test cannot correlate with another separate test any higher than it can correlate with itself, i.e. estimated subscale intercorrelations will always be lower than the true value for the subscale intercorrelation, because of what psychometrists call 'attenuation due to unreliability'. So given that generally the test-retest reliabilities for Tobacyk and Milford's original subscales fell somewhat short of being satisfactory (only Psi Beliefs and Extraordinary Life Forms were above Kline's (1993) accepted minimum of .80) we should note a particular degree of caution in dismissing evidence for relatedness in the subscales. In fact, if one calculates corrected Pearson correlations (using the formula for correction of attenuation due to unreliability given on page 14 of Kline, 1993) then we see that, for example, the revised estimate of intercorrelation between Psi Beliefs and Spiritualism lies at around the .66 level, and that for Psi Beliefs and Precognition is around .56. These figures certainly cannot be dismissed lightly, though I would advise caution in placing too strong a faith on an estimate of an estimate! Nevertheless, the original correlations are sufficient evidence for relatedness between factors for most psychometricians.

Perhaps the most revealing evidence for lack of orthogonality in the revised PBS is the fact that Tobacyk (1991) reports significant correlations between the two largest factors of his Prediction of Future Events Scale and the Superstition subscale of his PBS. As the reader will be aware, Tobacyk (1988)

used the two largest factors of his Prediction of Future Events Scale to form the new questions for his PBS Precognition subscale. Clearly we must regard the orthogonality of these two subscales with some suspicion. Unfortunately, Tobacyk (1988) provides no new evidence to suggest that the revised PBS contains truly orthogonal factors. Presumably orthogonality was assumed from the previous analysis of the old PBS. But as the discussion in the first half of this section shows, the evidence for orthogonality in that scale was based on a too premature dismissal of the (in some cases quite strong) evidence for correlations between subscales.

Evidence for construct validity assessed

Much of the above discussion of the subscales can be regarded as an in depth examination of the general validity of the PBS subscales, and we have seen that there is cause to challenge both the face validity (e.g. the item 'There is life on other planets' does not even look like it taps into UFO type beliefs), and construct validity (e.g. both Psi beliefs and Precognition are poor measures of their respective constructs, for reasons of comprehensive coverage and item content respectively) of these subscales. However, Tobacyk and Milford (1983) devote a substantial portion of their paper to demonstrating the construct validity of the PBS and its subscales empirically. Construct validity was first introduced into the psychometric literature by Cronbach and Meehl (1955) in an attempt to get around the difficulties in establishing certain specific types of validity for certain sorts of test. Construct validity refers to whether the concept or construct being measured by the scale is in reality what our scale measures. Evidence for construct validity may subsume evidence for all other kinds of validity (concurrent, predictive, face, incremental and differential) and may be regarded as the most general sort of validity that one can demonstrate. Thus to show evidence for construct validity one needs to show quite generally that the empirical evidence derived from use of one's scale accords well with the definition of the psychological trait one has proposed.

Tobacyk and Milford (1983) established the construct validity of their scale by collecting together seven other measures of psychological traits; the Internal-External Locus of Control Scale (Rotter, 1966), the Sensation Seeking Scale (Zuckerman, 1971), the Threat Index (Hays, 1974), a self-esteem measure (Hays, 1974), the Uncritical Inferences Test (Haney, 1954), the Dogmatism Scale (Rokeach, 1960), and an Irrational Beliefs Questionnaire (Newmark, Frerking, Cook and Newmark, 1973), and go on to correlate these measures with the global score and subscales for paranormal belief on the grounds that, "...due to theoretical considerations, they were hypothesized to be related to paranormal beliefs." (pp. 1034).

Tobacyk and Milford report correlations for these variables with the PBS ranging from 0-.34, and argue that these results broadly support the construct validity of the PBS. I could at this point discuss in detail the discussions of the evidence for construct validity given by Tobacyk and Milford, but I shall not do this for a very simple and basic reason; *none of the evidence reported by Tobacyk and Milford for construct validity is in fact evidence for such validity.* Instead I want to explain why in fact such evidence cannot *in principle* be evidence for the construct validity of the PBS.

To establish the construct validity of a paranormal belief scale one needs to show that it is feasible to define a construct - 'paranormal beliefs' - and that the behaviour of persons (both in surveys and in life generally) is consonant with such a construct. So to show construct validity for the PBS one needs to establish some basic facts. Firstly, one needs to show concurrent validity, i.e. that this measure of paranormal belief correlates reasonably well with previous measures of paranormal belief (assuming of course these previous measures are valid). Secondly one needs to show that the PBS can account for differences in persons that are essential and intrinsic to the notion of paranormal belief. Thus, it must be essential to any concept of paranormal belief that to have such beliefs one must have had some acquaintance with the topic of the paranormal, from whatever source (experience being the most obvious). So, if

our PBS could not account for differences in level of acquaintance with the paranormal (which could take the form of differential or predictive validity) then we would need to seriously question the validity of the scale. For example, if we measure the degree of paranormal experience one has had we would expect people with high experience to show high paranormal belief, and vice versa. Any evidence to the contrary would suggest that our scale was not measuring paranormal beliefs. And, of course, we could provide further evidence of this type for such things as reading books on the paranormal, visiting mediums, psychics, astrologers for advice, and taking seriously such advice as given. Even more basically, we would need to predict such things as that measures of paranormal belief taken now will continue to predict levels of paranormal belief in the future, and that our measure of paranormal belief does not correlate strongly with measures of things that are not intrinsic to paranormal belief (i.e. are not measures of paranormal beliefs). Thus, we would not expect a measure of paranormal beliefs to correlate strongly with a measure of attitudes towards normal things (e.g. smoking, party politics, tax increases).

It should now be quite obvious that the 'evidence' for construct validity offered up by Tobacyk and Milford is simply not evidence for validity at all. The fact is that Tobacyk and Milford's analyses *presume* the validity of the PBS. None of the scales used relate essentially or intrinsically to the concept of paranormal beliefs (or events associated with paranormal phenomena). Indeed, the scales that are used to support the validity of the PBS can only be used if it is a defining characteristic of such beliefs (i.e. it is a part of the construct) that they are intrinsically dogmatic, uncritical, irrational, sensation seeking, death threat oriented, governed by a concern with external determinants of behaviour, and are bound up with self esteem. That in general parlance and in Tobacyk and Milford's definition of the paranormal none of these characteristics appear as intrinsic qualities is clear evidence that these things cannot be brought to bear on the evidence for the construct validity of the PBS. Of course these things may certainly relate to paranormal beliefs for reasons which are of theoretical

interest, but to establish such relations requires that one has a valid measure in the first place. Thus, Tobacyk and Milford (1983) provide no good evidence that the PBS measures what it claims to measure - accepting some evidence for face validity of some of the items - and one can make this claim even though I do not doubt that had such evidence been looked for, it would have been found to some extent.

To sum up this section we may reasonably conclude that beyond some evidence for face validity and some content validity, the actual evidence for the validity of the PBS is very poor. Subsequent evidence from other researchers has shown some concurrent validity for the PBS (i.e. it correlates with other measures of paranormal belief; Thalbourne has several unpublished studies reporting high correlations between the PBS and his sheep-goat scale, and see Irwin, 1985) but this type of validity is on its own not sufficient to recommend the use of the scale.

Problems of poor scale use

At times researchers have not used the scale in the spirit for which it was designed. The revised PBS presents seven independent subscales of paranormal belief for the researcher to use as such. The primary rationale in constructing the scale was that paranormal belief is multi-faceted, and not just amenable to analysis in terms of a general factor of Paranormal Belief (scales that have treated belief in the paranormal as an all or none property have tended to seem too simple and superficial, e.g. Scheidt, 1973, and Randall and Desrosiers, 1980). Nevertheless in an important and highly consequential area like the study of reasoning and belief in the paranormal we can observe Wierzbicki (1985) correlating subjects' scores on a syllogistic reasoning task with the global score for paranormal belief and nothing else (see Irwin, 1991a, for a credible attempt to approach the same issue, and then again Irwin, 1992, correlating childhood trauma solely with the full scale PBS score).

There is no particular problem with reporting full scale score correlations from a multi-factorial scale (this is often done) when the meaning of the full scale score is clear. However, I would argue that the meaning that one should give to a full scale score on the PBS is not as clear as it might seem. The full scale score may certainly be tapping into general paranormal belief. But let us consider the sort of person who might score highly on general paranormal belief as assessed by the PBS. This person would necessarily have expressed strong belief in so many diverse phenomena, varying widely in exotic appeal, that one might just as easily be tapping into mental health and gullibility factors too. This is especially so given that the PBS was designed to assess as wide a range of phenomena as possible. But the real problem here stems from the expansive definition of the paranormal. This leads us directly to a discussion of the definition of 'paranormal'. Indeed, in this discussion we may get to the heart of the problems concerning the construct of "paranormal beliefs".

Problems in defining "paranormal"

Lastly, I would like to raise concerns about Tobacyk's (1988) definition of the paranormal. In the revised PBS only Broad's (1953) definition of paranormality- 'explicability achieved only by major revision of the basic limiting principles of science'- is used. The problem with this definition is that it does not seem to fit the scale items adequately. Thus, with respect to Big Foot, Nessie, or extra-terrestrials, surely the existence (or not) of any of these things is merely a matter of natural historical fact. Only detection of them with certainty will show that they do exist, but their paranormal status would have to be much the same as that of the Coelacanth fish discovered off the coast of Southern Africa in the 1930's. Until its discovery the fish was not in principle ruled out of all existence, just considered merely unlikely. But, as a matter of natural historical fact, it still existed and was later discovered.

But Tobacyk and Milford (1983) do acknowledge that their definition of paranormal is problematic. Indeed, Tobacyk notes (pp. 7), "....Extraordinary Life Forms...does not strictly conform to the "basic limiting principles"

definition of paranormality." It seems as though Broad's definition of paranormal in terms of "basic limiting principles" is not as clear cut as it might seem. As Braude (1978) points out, "Broad has failed to explain what, in general, a phenomenon must conflict with in order to conflict with a B[asic] L[imiting] P[rinciple]." (pp. 251). Clearly a better definition is needed.

I propose that we use the far more restrictive definition of paranormal offered by Irwin (1993, he attributes it to Thalbourne, 1982, but this is not in fact the case). Irwin defines a thing as paranormal if it "Refers to hypothesised *processes* that in principle are 'physically impossible' or outside the realm of *human* capabilities as *presently conceived by conventional scientists*" (pp. 1) [emphasis added]. On the grounds of this new definition Witchcraft, Psi Powers, Precognition and Spiritualism would be retained as factors. They clearly refer to human capabilities and hypothesised processes and all of them share a certain fundamental relationship: the dependence on unexplained principles of nature (Psi processes apply to Psi, Precognition and Spiritualism - Magical Energies to Witchcraft).

Conclusions

From what I have outlined in this chapter we can only conclude that one of psychology's most widely used scales for the assessment of belief in the paranormal, the PBS and revised PBS, is less than satisfactory for its intended purpose. I have outlined what I believe are serious objections to the view that paranormal belief can be accounted for by seven dimensions. These objections are based in part on problems with selection of questionnaire items, and for the most part result from the all-encompassing factor selection procedure used in the original factor analysis. Quite simply, according to a more accurate factor selection criteria there are not seven separate factors of paranormal belief, there are four; Traditional Religious Belief, Psi Belief, Witchcraft and Superstition. Further still, whilst there are most certainly less than seven significant factors in the PBS, those that would remain on more accurate

criteria would most likely not be orthogonal (based on a perusal of evidence reported by Tobacyk and Milford, 1983, and Tobacyk ,1991).

Clearly we need a new PBS. Much useful and insightful research has been produced with the old PBS and its revision, but the problems present in those scales necessitate a new scale, adequately designed, analysed and implemented. Only with this can a firm foundation be made for further research in this area, and as there are no signs that research on the topic of paranormal belief (in either psychology or parapsychology) is slowing up, we can only hope to see a new PBS in the not so distant future.

However, before we begin to construct a new PBS I want to note a number of concerns that may influence the exact nature of such a scale. My first concern is that our new PBS, suitably narrowed in content to contain less that is merely strange and wacky and more that is in line with the kind of humanly relevant definition of paranormal phenomena that I advocate, might run the risk of containing subscales (say of ESP, PK, and Life after Death) that contain little more than a series of bloated specifics. Catell (1973) has shown that scales with excellent reliabilities can be obtained merely by filling each scale with items that do nothing but paraphrase each other. Thus, the two items 'ESP allows people to know each other's thoughts' and 'People use ESP to read other's thoughts' are little more than minor modifications on the same theme. They would correlate highly with each other, and give good internal consistency and test-retest reliability, but either question alone would work just as well! My suspicion is that it is all too easy to construct face valid and reliable belief scales using bloated specifics. This worry is compounded by a related piece of evidence, taken from the meta-analysis of the sheep-goat effect that I have previously reported (Lawrence, 1993). In that meta-analysis, Schmeidler's (1943) single question measure of belief 'Do you believe ESP can occur under the conditions of this experiment?' gave effect size estimates for ESP success better than almost every multi-item measure of belief in the paranormal. Clearly some of this success can be attributed to the enhanced consistency of attitude-

behaviour relations found when highly specific measures of attitude are used (and it's fair to say that Schmeidler's question puts the subject on the spot). But the implication may also be that you can tap in to belief just as well with one, two or three questions as you can with 25 or 50.

There are two solutions to the problem of the bloated specific scale if it exists (we need to be sure that the problem exists by comparing, say, a six item measure of paranormal belief; measuring clairvoyance, precognition, telepathy, PK, life after death and magic, with a full blown multi-item multi-factorial scale. The present PBS could be used for this purpose. If they give rise to similar results then bloated specifics are a problem). Firstly, we could stick with the smaller, more convenient scale. Secondly, we could try to produce a genuine multi-factorial scale which does not contain bloated specifics. One way to do this would be to index paranormal beliefs through behavioural indicators coupled with more belief like statements, akin to the way that extroversion is measured. However, in doing this we should be aware of yet a further difficulty; to what extent can we treat paranormal belief like a basic personality trait at all? Clearly paranormal belief is more akin to a socially held attitude, developed through paranormal experience and socialisation by the wider culture. In short, whereas to measure someone's level of extroversion is to assess directly certain behaviours that are argued to result from a basic biological/social propensity for such behaviour essential to that person, to measure someone's level of paranormal belief is to measure a number of beliefs held upon evidence that is diverse socially, culturally and psychologically.

There is a third option however. In the third option we retain the benefits of the multi-factorial scale but change the function of the scale from simply measuring beliefs to measuring beliefs in association with their most obvious and immediate social and phenomenological causes. This third option does real justice to the aim of measuring belief as socio-culturally and experientially based, whilst actually enhancing our capacity to rationally dissect the results of pure paranormal belief research! How does one do this?

Let us consider the basic causes of paranormal belief. Firstly and obviously there is personal experience of ostensibly paranormal phenomena. Secondly, one may be acquainted with relatives or close friends who have had paranormal experiences. Thirdly, one may have experienced in the media popular/serious, academic/occult, scientific/religious works relating to paranormal phenomena. Lastly, one may not particularly have had or sought out any of the previous things, but instead base one's beliefs on a general world view, provided by one's education. These four basic areas then; personal, interpersonal, socio-cultural and educational, are the most basic immediate causes of one's paranormal beliefs. Note the diversity in sources of belief. Note that to have based one's beliefs on influences from any one of these four does not necessarily lead to the conclusion that one has influences from any of the other three causes (in principle the four causes are independent). In practice, these four basic influences will be moderately related.

At present all paranormal belief researchers use measurement devices relating solely to the beliefs themselves. A most common procedure in research with these scales is to correlate paranormal belief with other psychological factors like paranormal experience in order to better understand the mind of the paranormal believer. This endeavour is I would argue confounded by one simple, but devastating flaw - someone who scores highly on a measure of paranormal belief could do so for any combination of influences from the basic four just outlined. Thus psychics score high on measures of paranormal belief, but so do parapsychologists and people who read about the paranormal in the Fortean Times. These various groups are a world apart in their influence from the four basic causes, yet on our measure of paranormal belief and (most perniciously) in our empirical data these diverse basic differences remain undifferentiated. This obviously has disastrous, albeit hidden, consequences for the development of understanding in paranormal belief research. Thus, our basic failure to acknowledge the diversity of types of believer leads us to develop multi-dimensional measures of belief which are at other levels more

strangely and subtly one dimensional! This point of course applies quite generally to any of the present measures of paranormal belief.

Given that this is a major problem, a solution becomes particularly pertinent. In short, I propose that we develop a measure more along the lines of a *paranormal beliefs and influences scale* (PBIS). This scale would certainly measure paranormal beliefs, but would also measure the weight of influence from our four basic causes of belief (I realise that I am making a presumption in supposing paranormal experience to be a cause of belief but evidence I have reported elsewhere favours this experiential source hypothesis over other more attributional hypotheses; see Lawrence *et al.*, 1995). Only then do we have a realistic chance of getting inside the mind of the sheep and goats (whatever *species* they might be!).

Such a scale is unlikely to be produced for some time yet. As we have seen, building a valid and reliable measure of anything is a tricky business. Hopefully, future efforts at scale construction can involve input from both parapsychologist and psychologist alike, both paranormalist and sceptic. This is indeed necessary if we are to break with the partisan and frankly irrational 'these are our scales and those are yours' distinction that has split scale use between psychologists and parapsychologists with respect to paranormal belief.

Nevertheless, present circumstances accepted, Thalbourne and Delin (1993) have most recently reported a revision of their Australian Sheep-Goat Scale that shows some promise in the assessment of the core concepts of paranormal belief (ESP, Psychokinesis and Life after Death). It remains for that scale or any other to prove its worth as the scale to use. If the PBS is to be even a contender for that distinction it will have to change. And if it does change, then the enterprise undertaken in this chapter will seem to me all the less negative, and all the more productive for it.

At this stage in the thesis I wish to direct the reader's attention to the appendix, where one may find two papers written by Tobacyk in response to my critique (as it appeared in the Journal of Parapsychology). The reader is advised to read Tobacyk's initial counter reply, before going on to read the next chapter, which forms my own discussion of his counter. Then one may read Tobacyk's final comments on my own reply, before we leave the critical component of discussions of the PBS, and move on to chapter 5, where I present the results of am empirical investigation of the factor structure of the PBS, in an attempt to settle our disagreement of the nature of the PBS.

Chapter 4

THE BATTLE IS JOINED!

TOBACYK'S DEFENSE OF THE PBS AND THE ENSUING DEBATE.[1]

In the March 1995 issue of the Journal of Parapsychology, after my own critique of the PBS, Tobacyk presented a spirited, thought provoking, and constructive reply to my criticisms (Lawrence, 1995a). His reply showed perfectly how critical dialogue can lead to significant and constructive suggestions for future research, and help to bring out problematic issues relevant to the enhancement of our understanding of complex areas of research. Also it became clear to me from his reply that there were many areas where Tobacyk and I are were in agreement. Quite generally, we seemed to agree that we needed an improvement to the PBS, and that the PBS has a number of problems which only future research would be able to clear up, though we differed over just how severe these problems are. More importantly I think, both our papers made substantive contributions on the topic of how new research in this area might proceed. In what follows I outline, with as great a degree of fidelity and impartiality as possible, the substance of Tobacyk's reply, and having done so deal primarily with those areas where we disagree, before concluding my methodological efforts with some final comments.

---

[1] The reader should acquaint theirself with Tobacyk's response to my critique, given in the appendix, before proceeding with this chapter.

Areas Where We Disagree

Problems of defining 'paranormal':

Although I agree with much that Tobacyk says about the problems of defining paranormal, Tobacyk picks up on a subtlety in Irwin's definition (the use of the word 'or' rather than 'and') and goes on to show that, therefore, Irwin's definition of paranormal is really not as restrictive as Broad's. If one sticks to the letter of Irwin's definition then Tobacyk has a point: 'paranormal' Irwin-style may be little different from 'paranormal' Broad-style. Given this subtlety, perhaps I need to give my own definition. Something is paranormal if it is 'in principle physically impossible *and* outside the range of human capabilities as presently conceived by conventional scientists'. I prefer this more restrictive definition not so much because I want it to restrict certain phenomena, but because I believe it is important for our research to be humanly relevant. I do not have space enough to go into detail however (instead I think the problem of defining the paranormal needs sustained critical discussion - that will need another paper!).

We both acknowledge that the extra-ordinary life forms subscale is not paranormal (it is, to paraphrase Tobacyk, paranormal by association). Instead, Tobacyk makes much of what I don't say about Traditional Religious Belief and Superstition. In fact, I say little about the paranormal status of traditional religious belief and superstition because my thoughts on the paranormal status of these beliefs are equivocal. I can see reasons why religious beliefs might be regarded as paranormal, but then again I see reasons why calling religious beliefs paranormal might be problematic (e.g. religion depends upon *socially* organised faith very strongly, paranormal beliefs depend much more on *personal*, human experience). As for superstition, there have been two ways of defining it in the past. Firstly, superstition has been defined, mostly by academics, in such a way that it is identical with paranormal belief (e.g. Planer, 1988). Secondly, superstition has been defined, mostly in folklore, as referring more specifically to certain events and occurrences which have certain socially

signified meanings (i.e. black cats are unlucky, one should throw salt over one's shoulder if one spills salt). If one adopts the first definition then all one needs to do is decide whether the PBS is the PBS or the SBS (Superstitious Beliefs Scale). This is hardly a substantive contribution. If one adopts the second definition then arguably superstition ceases to be paranormal. To say that black cats are unlucky is not to say that humans are possessed of some unknown capability to be unlucky in the presence of black cats. To conclude, I think Tobacyk's reply establishes the need to discuss problems of definition as an integral part of empirical research, and not as a separate component of purely philosophical or conceptual interest.

Adequacy of the PBS Factor Analysis:

My discussion of the adequacy of the PBS factor analytic evidence essentially centres around three things; (i) problems of factor selection, (ii) problems of factor rotation, and (iii) evidence for fewer factors from my scree slope analysis.

(i) Problems of factor selection - On problems of factor selection maybe Tobacyk and I are just going to have to differ over Kaiser-1. He can give studies that attest to its accuracy in some circumstances, and I can give studies which attest to the opposite. Thus, Kline (1993) notes that,

> "Catell (1978) has shown that rotation of too few factors tends to produce second order factors at the first order. Rotation of too many factors causes factors to split. A common solution to this problem and the default solution on many computer packages is to rotate all factors with eigenvalues or latent roots greater than one. However Catell (1978) has shown that with large matrices this is an overestimate of the number of significant factors and Cliff (1988) has suggested that this criterion be abandoned. Barrett and Kline (1982) examined a number of methods of selecting the correct number of factors and showed that

two methods appeared to reach the best solution ...These were the Scree Test (Catell, 1966) and the Velicer method (Velicer, 1976)." (pp. 122)

At least one needs to be wary of the Kaiser-1 criterion, but given the fact that Tobacyk uses three other criteria I think it quite fair to note caution over the stability of the factor solution. Thus, with four criteria for factor selection, does a factor have to satisfy all four criteria, or just three, or does one chop and choose criteria in order to get the factors that one desires? I think these are significant issues that may have played a role in determining the seven factor solution.

Also, I don't think that a factor solution that accounts for 60.7% of the variance with 13 factors can be seen as satisfactory. Quite generally, satisfactory solutions, even taking into account error ridden social and psychological data, tend to be more in the region of 60-70% variance with 4-5 factors. Factor analysis is essentially about reducing complex sets of observed variables to their essential latent causes. When one needs 13 factors to achieve 60% accountability, or 7 factors to achieve 48% accountability (which is actually the case with the PBS) then this is hardly a victory for parsimony. Rather, it suggests the researcher has brought together a wide and disparate set of unrelated variables, of little general explanatory consequence.

(ii) Problems of factor rotation - As my comments about the orthogonality of the PBS show, the PBS does contain a core of beliefs which do correlate together; strangely, precisely those factors which meet the human capabilities definition of paranormal. Note that Tobacyk simply doesn't address my concerns over subscale orthogonality however, instead he simply reasserts that promax oblique rotations showed up little evidence for correlated factors. At the least I would suggest future researchers look strongly at the evidence for oblique factors in paranormal beliefs research. Consequently, I maintain that the

problems of factor selection and orthogonality/rotation in the PBS remain.

(iii) Problems of the scree plot analysis - As Tobacyk points out in his reply, 'Lawrence's use of the scree plot analysis to determine the meaningfulness of rotated factors is unconventional and probably invalid' (pp. 6). In fact, the scree slope analysis, if performed on rotated factors, is definitely invalid. The reason for this is as follows; as Tobacyk himself notes, the scree plot analysis is used to determine the number of factors for rotation, by plotting a visually judged line of best fit through the smaller factors (expressed as eigenvalues) of the initial unrotated factor solution. Although subsequent rotation of factors does not change the overall amount of variance explained by the factor solution, it does change the specific amount of variance accounted for by each individual factor. Thus, in order for my scree slope to be valid, we *must* be sure we have the initial solution eigenvalues. On the face of it I appear to have made a mistake. However, two points were made to me by a colleague who specialises in psychometrics that set me on a quest, the resolution of which proves highly revealing. The two 'problems' pointed out to me are as follows[2];

(i) Given Tobacyk's use of principal factor analysis with Varimax rotation, the distribution of his eigenvalues is highly unlikely if they are genuinely from the rotated solution.

Thus when we look at Tobacyk's plotted eigenvalues we see that there are some marked factors which account for a great deal of variance, and many other factors which account for a very small amount. The problem is that Varimax rotation very rarely produces this kind of factor distribution. Indeed, varimax rotations by definition tend to maximise, and hence equalise the variance

[2] Thanks to Dr. Ian Deary, Professor of Psychology at the University of Edinburgh.

accounted for by factors in the rotated solution. In short, Tobacyk's data look much more like they come from the initial solution, and could only really have come from the rotated solution if Tobacyk and Milford had markedly different numbers of items for each factor.

(ii) Although the published data from Tobacyk and Milford (1983) are given in percentages of variance accounted for, the list of factors sent me by Tobacyk clearly give eigenvalues alone. But eigenvalues are of no consequence in the rotated solution, and many factor analysis packages do not give them for the rotated solution.

Now, these two pieces of evidence clearly suggest that Tobacyk has sent me the details of his *initial* solution results. We can confirm this by finding out whether Tobacyk's statistical package gave eigenvalues for the rotated solution. If the Stats package used by Tobacyk did not give eigenvalues in the rotated solution then the validity of my analysis stands. Tobacyk used a pre-1983 version of SAS to perform his factor analysis, so I contacted the SAS Institute in Britain to ask if that version gave eigenvalues for the rotated solution. SAS were very helpful, and Mark Dawson, technical support advisor, contacted SAS in the US where details of earlier versions were kept.

SAS US still had a mainframe running a pre-1983 version of SAS(!), so they ran a principal factor analysis with Varimax rotation on some old sample data, and the results were clear. *SAS pre-1983 does not give eigenvalues for the rotated solution.* Thus, from the two points made above, and the confirmation from SAS, we can only conclude that Tobacyk has unwittingly sent me the eigenvalues for the initial solution, and would appear to have used them also in his published report. I should perhaps add that it is not at all uncommon to find researchers reporting initial solution eigenvalues as though they were the rotated solution

results (indeed, I gather that with SPSS the way data are presented often leads to this error).

The above discovery validates my use of the scree slope analysis. However, Tobacyk argues that even if the analysis were valid, the accuracy of the slope is dubious. Thus, Tobacyk suggests that the data from factors 14 to 25 are needed to really do the scree plot. This statement reveals two basic errors however, thus;

(1) Tobacyk forgets that he performed *principal factor analysis*, and not *principal component analysis,* so 13 factors is in fact how many factors need to be analysed by the scree slope method, and not 25.

(2) Indeed, even if principal component analysis had been used Tobacyk is wrong in saying that there would be 25 factors, as principal components extracts as many factors as variables, which in Tobacyk's original unfactored sample was 61.

In any case, Tobacyk's claim that we need those factors to identify the departure from linearity (i.e. random factors) is not appropriate. Thus if the smaller factors were essentially random (as is usual and as Tobacyk seems to indicate) then they would quite generally show a linear increase in value from close to zero to close to one, in which case approximating this rough linearity with a smooth line will not give great inaccuracy. But even if my estimation of the plot of unknown factors is inaccurate, how much room for a radical departure from linearity is there between an eigenvalue of just above 1 (for factor eight) and exactly one (for factor thirteen)? It is in fact impossible for there to be a jump in the smaller factors larger than the jump that is clearly evident in the larger factors (from 2.31 for Superstition to 4.09 for Witchcraft - an increase in accountability larger almost than the whole range of accountability possible in the factors from Spiritualism to factor 13).

Indeed, one could even argue that my placing of the line is *conservative*, because the rationale of the scree slope is to identify a marked increase in variance explained sufficient to show that the 'mountain' of genuine factors 'starts here' - all the rest being mere rubble and debris. The first truly marked deviation from linearity is represented by the jump from Superstition to Witchcraft, and placing a line to reflect this gives only three factors! Thus I think that the scree slope analysis is both valid and largely accurate (any error is unlikely to make a difference). But we could argue over small details *ad infinitum*, so is there any new evidence for the factorial structure of the PBS?

I have details of at least one attempt to factor analyse the PBS, performed recently by the Dutch researchers Tomic, van der Sijde, and Snel (unpublished), in which they found only five factors. Of Tobacyk and Milford's seven factors, only two were reproduced without change in the Dutch study; Traditional Religious Belief (their strongest factor) and Witchcraft (their third strongest). Of particular interest is that the Dutch Psi scale consumed Tobacyk's Psi, Spiritualism, and Precognition subscales, and Superstition and Extra-ordinary lifeforms loaded onto the same factor (which the Dutch group termed 'Superstition'). The last factor, which the Dutch group termed 'Impossible' is a strange little factor composed of items that were negatively phrased. In general then, the Dutch data, based on a moderately large sample of over 300 Dutch men and women, supports fewer than seven factors. Interestingly, the Dutch Psi subscale shows a clear degree of relatedness between just those subscales of the PBS which come under the human capabilities definition of paranormal.

In short, the problems of factor selection and orthogonality remain, and on deeper analysis my scree slope analysis appears to be valid also. When we incorporate new *empirical evidence* bearing on the number of factors problem, we once again find fewer than seven factors.

The correct dimensionality of paranormal beliefs:

In a somewhat paradoxical manoeuvre Tobacyk resorts to relativistic/constructionist arguments to challenge the notion that one can determine the correct dimensionality of paranormal beliefs. This attempt to bypass the substance of my critique is however thoroughly unsatisfactory. Thus, if we take the relativist position seriously then we simply cannot legitimately ask questions about the factorial structure of paranormal beliefs at all. But Tobacyk's efforts surely presuppose that factor analysis has *some* value for research purposes.

Perhaps Tobacyk means to advocate a weak relativist position whereby cultures, people, and societies will simply differ in some or other respect on the PBS. But this weaker position turns out to express little more than trivia; we all know that people and cultures differ, the question is whether there is some invariant factor structure of beliefs within which we can express these differences. The PBS is highly problematic when viewed from this non-relativistic perspective (e.g. how can African natives be expected to express agreement with an item on the Loch Ness monster?). I feel, in contrast to Tobacyk, that we can ask legitimate questions about the general structure of paranormal beliefs, and that it is a challenge to devise such scales in a way which makes their underlying factors culturally invariant, even if average level of belief on the scales differs across cultures.

The reasonableness of my critique of the PBS:

This section of Tobacyk's reply deals largely with my comment on the extra-ordinary life forms and precognition subscales. I believe I've said about as much as I need say about the extra-ordinary life forms subscale in my original paper; thus, (a) ELF is strictly not paranormal (as Tobacyk agrees), and (b) it is clear that some of the items in that scale are far from adequate.

As for Tobacyk's lengthy defence of the PBS precognition section, I would say it is excellent save for one thing; unfortunately, all the points made by Tobacyk only really bite if his precognition subscale actually measures precognition. As I've already noted in my earlier paper, the precognition subscale of the revised PBS barely measures precognition. Firstly, the precognition subscale is composed of two factors one of which is totally unrelated to the notion of precognitive prediction of the future whilst the other contains only one definite item on precognition. Secondly, and worse still, the two factors which make up the 'factor' of precognition are unrelated to each other! Consequently, anything Tobacyk says about the validity of precognition in the revised PBS needs to be weighted against the fact that his revised precognition subscale barely measures precognition, and anything he says about the validity of the old precognition subscale needs to be weighted against Tobacyk's own admission that the items in that subscale were ambiguous enough to warrant a new scale. Under these circumstances it is highly debatable whether one can say anything of consequence about the validity of the precognition subscale.

The construct validity of the PBS:

Tobacyk doubts my assertion that his evidence for construct validity of the PBS is in principle not such evidence, and asserts that in any case the PBS has 'strong face and content validity'. I do not doubt that the items on the PBS are face valid, although I am sure some of the items on the PBS have *no* content validity (i.e. the precognition items of the new PBS, or the Extra-ordinary life form item). Even if the PBS did have strong face and content validity, I do not think this is sufficient to recommend its widespread use. Ultimately, construct validity is required, and I maintain that Tobacyk's initial evidence for this was no such evidence at all. Also, I disagree that construct validity is established through the construction of a rich nomological network of interrelations with other non-paranormal factors. In so far as Tobacyk genuinely intends this to mean that the PBS interacts with other types of scales

which do not themselves refer to the concept of the paranormal, then the reasoning behind such evidence could only *presume* construct validity, and not *establish* it. Thus, if we establish that the PBS correlates with gullibility, as some might predict from their construct of paranormal belief, this may only establish that our PBS is an obscure indicator of gullibility.

Nevertheless, Tobacyk suggests that evidence for the construct validity of the PBS is present, and cites one example (Tobacyk and Mitchell, 1987). I think this evidence is instructive in that it does establish some construct validity, but not of the sort that Tobacyk's construct of paranormal would require (whatever that might be). Instead, Tobacyk's evidence for construct validity supports a notion of paranormality with which we should now be quite familiar - namely, the view that a cluster of paranormal beliefs are evident which take as their focus the fact that they all refer to '*hypothesized human capabilities or processes*'! Just take a look. They *all* fit this definition. Obliqueness rears its ugly head again. But given the remarks of Beloff (1963, p. 111) we should perhaps not be surprised at the human focus, thus, '...paranormal phenomena...almost invariably occur in connection with a human person. Now there is no a priori reason why this should be the case. One could conceive of a world in which inanimate objects took an occasional paranormal holiday during which they behaved entirely capriciously.' Thus, the human connection seems to be a *central defining element* in paranormal phenomena, of just the sort that a solid construct of the paranormal could be built on.

Final Comments

Despite our disagreement over the substance of my critique, Tobacyk makes many valuable points. Thus, Tobacyk has many good suggestions to make about how future research might proceed. My own recent efforts to improve the methodological and theoretical rigour of paranormal beliefs research echo many of the points raised by Tobacyk (see Lawrence *et al.*, 1995). Perhaps most

importantly however, is the fact that Tobacyk clearly sees a greater role for understanding paranormal beliefs than merely enhancing knowledge. Indeed, in his parting statement emphasising the fundamental human questions and the relevance of paranormal experiences/beliefs research in answering these questions, Tobacyk displays a sensitivity to the wider issues sadly lacking in much contemporary parapsychology. In this respect Tobacyk and I share the same basic attitude. It is precisely because these issues are important that we must devote sustained critical attention to the improvement of our scales for the measurement of paranormal belief (and indeed, of any variable). What is more, lest we waste time improving scales which are beyond improvement, we must focus our critical efforts on those scales which do show promise.

I wish to thank Tobacyk for the dignified, temperate, and productive nature of his reply. May we soon see someone produce a new PBS, rising like the proverbial Phoenix from the ashes of the old!

Chapter 5

CONFIRMING THE FACTOR STRUCTURE OF THE PARANORMAL BELIEFS SCALE: BIG ORTHOGONAL SEVEN OR OBLIQUE FIVE?[1]

As we saw in the last chapter, despite its widespread use in research into the psychology of paranormal belief, the Paranormal Belief Scale (PBS) (Tobacyk and Milford, 1983; Tobacyk, 1988) has more recently come under criticism on statistical, methodological and conceptual fronts (see also Lawrence, 1995a; and the subsequent replies from Tobacyk, 1995a; and Lawrence, 1995b). The essence of my criticism was that the original evidence for seven independent factors of paranormal belief was flawed, and that in addition there were serious problems with the scale's validity, sufficient to warrant construction of a new scale. I argued that, in contrast to Tobacyk and Milford's seven factor solution, a more realistic description of the factor structure of paranormal belief was a four factor solution with oblique (i.e. correlated) factors. Such a view can be justified on two separate grounds; statistical and conceptual.

Statistically, I argued that Tobacyk's four criteria for selecting significant factors were too liberal, and hence prone to over-extraction. If one took these considerations into account, and applied a more conservative criterion for selecting factors, then four factors are suggested. In addition, I argued that evidence for correlations between factors had been overlooked, and that consequently the subscales of the PBS were unlikely to be uncorrelated.

---

[1] The first study in this chapter was carried out as part of a wider study conducted in collaboration with my colleagues Dr. Chris Roe, who collected the data, and Dr. Carl Williams, who had independent variables included in the study unrelated to this thesis. All the reported analyses and interpretations are wholly mine. The first study was subsequently published in the Journal of Parapsychology, see (Lawrence, Roe and Williams, 1997). The second study was carried out in collaboration with Pio De Cicco, my project student at Coventry University. I extend my thanks to him for the hard work he put in collecting the data.

Conceptually, I challenged the paranormality of the Extraordinary Lifeforms scale, and expressed doubts about the definition of paranormality used by Tobacyk (1988), suggesting a new definition of paranormality based upon that of Irwin (1993). With such a new definition there would be four definite factors of paranormal belief - Psi, Spiritualism, Precognition, and Witchcraft; the paranormality of traditional religious belief and superstition being a matter of some dispute.

Whilst my critique (Lawrence, 1995a+b) served to establish some serious difficulties and uncertainties with respect to the factor structure of the present PBS, I was unable to come to any firm empirically based conclusion about the veracity of Tobacyk's 'orthogonal seven' or my suggested 'oblique four', primarily because Tobacyk's original data were unavailable for reanalysis. Indeed, a common conclusion reached as a result of the debate between Tobacyk and myself was the need for future research to use more confirmatory techniques to resolve speculations about the factor structure of the PBS. In this chapter I do just that, using the technique of confirmatory factor analysis using EQS (Bentler, 1989).

Confirmatory factor analysis (CFA) is a particular method of factor analysis, in which the researcher's aim is to confirm some hypothesised factor structure, given a set of correlational data. Thus, a researcher may suspect that the correlations between a set of measured variables can be accounted for by a three factor structure. In this case the researcher might suppose that measured variables one to ten are all indicators of factor one whilst variables 11 to 20 are indicators of factor two, and 21 to 30 indicators of factor three. In addition the researcher may suspect that certain pairs of factors in one's study are correlated, whilst certain other factors are not. Thus in my three factor example the researcher might suppose that factors one and two, and two and three are correlated, but not one and three. Also the researcher may suspect that some of the variables are valid indicators of more than one factor.

CFA permits the researcher to test the appropriateness of these kinds of assumptions for the observed data by letting the researcher impose meaningful constraints upon the factor analytic process. Typically this is done in two practical steps. In the first step, the researcher specifies exactly which measured variables are indicators of which specific underlying factors (in statistical language the researcher is said to be specifying the 'measurement model'). Once the factors have been identified with certain variables, the researcher goes on to specify which of the factors are correlated with each other, and which are not. This is in stark contrast to exploratory factor analysis which allows the researcher almost no freedom to specify the factor analytic process. Instead, exploratory factor analysis assumes (usually incorrectly) that all variables are indicators of all factors, and that all factors are either orthogonal (i.e. uncorrelated) or oblique (i.e. correlated). In addition, in order to achieve any solution at all, principal factor analysis has to make use of a number of generally arbitrary principles. Without these principles it would not be possible to arrive at any determined factor solution (this is the so-called factor indeterminacy problem). Indeed, Long (1994b) notes that, 'The exploratory factor model's inability to incorporate substantively meaningful constraints, and its necessary imposition of substantively meaningless constraints, has earned it the scornful label of garbage in/garbage out model.' (pp. 252).

In this chapter I am not especially concerned with finding out the factor structure of the PBS *anew*. Rather, Tobacyk and Milford (1983) have claimed (on the basis of their original exploratory factor analytic results) that the structure of the PBS is best represented by seven orthogonal factors (I shall refer to this model henceforth as the 'Big Orthogonal Seven'). In contrast, I very much suspect that the factor structure of the PBS is best described by a smaller number of factors, some of which may well be correlated. In the remainder of the introduction I outline my reasons for developing this suspicion as working hypothesis, and show how this working hypothesis can be

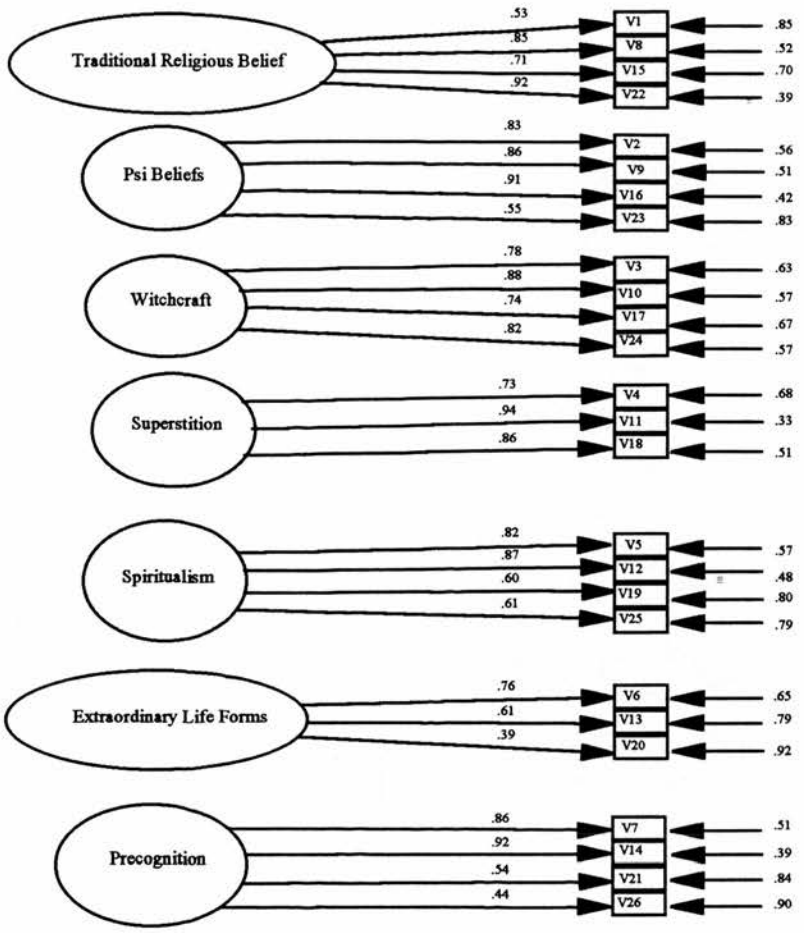

Figure 1. The Big Orthogonal Seven.

specified in terms of a confirmatory factor model postulating five oblique factors (my 'Oblique Five').

How the Big Orthogonal Seven are really the Oblique Five

Before I go on to show how, in the PBS, seven is really five and orthogonal is really oblique, I wish to briefly clear up some potential sources of confusion. Firstly, as I noted in the beginning of this chapter, there are both statistical/methodological and conceptual reasons to suppose that there are fewer than seven factors in the PBS, and that these factors are correlated. In this chapter however I am concerned only with implementing recommendations based upon my statistical/methodological critique of the

PBS. The reason for this is fairly straightforward: if one attempted to build a CFA model of the PBS based on my conceptual critique of that scale then one would begin by omitting the Extraordinary Lifeforms subscale altogether, and one may well decide against modelling Traditional Religious Belief and Superstition as well! Although such a model may well be preferable on conceptual and (ultimately) empirical grounds, a test of such a model would not provide a fair comparison with Tobacyk and Milford's (1983) Big Orthogonal Seven. Instead, one needs to show that even with Tobacyk's (1988) original subscale items all present there is still more support for the Oblique Five than for the Big Orthogonal Seven. In any case, as we shall see, the empirical resolution of my statistically motivated modelling attempts will lend weight to the conceptual model. Now I discuss my attempt to model the structure of the PBS proper.

In the previous chapter I identified a number of reasons why the factor structure of the PBS might be smaller than that found by Tobacyk and Milford (1983);

Firstly, I argued that Tobacyk and Milford's criteria for factor selection resulted in the extraction, and subsequent rotation, of too many factors. I suggested (somewhat speculatively) that perhaps four factors were more appropriate. Had four factors been extracted the resulting solution would most certainly have aggregated certain subscales, in particular those with higher correlations, such as the Spiritualism, Precognition and Psi Beliefs subscales, all of which come under the general rubric of 'psychic beliefs'. Thus, I chose to model all the Spiritualism subscale items of the PBS (5, 12, 19 and 25) and two of the Precognition subscale items of the PBS (21 and 26 -see below) as loading on a general factor of Psychic Beliefs.

Secondly, methodological problems with the development of the Precognition subscale almost certainly throw into doubt the status of that scale. Of particular

importance is the fact that the revised Precognition subscale is composed of two orthogonal factors (Psychically Gifted Persons and Paranormal Divinatory Systems). Thus at the very least one has reason to suspect that a better model would be obtained by dismantling the Precognition subscale to take cognisance of the fact that it is not unifactorial. One is left, however, with the question of where the two sets of items representing these factors should go. The items referring to psychically gifted persons (21 and 26) are clearly relevant to the Psi Beliefs subscale and in this chapter I model items 21 and 26 as loading on a wider Psychic Beliefs factor. Items (7 and 14) refer to astrological prediction rather than psychic prediction and consequently I model these items as loading on a factor of Anomalous Natural Phenomena along with the Extraordinary Lifeform subscale items (6, 13, and 20).

Thirdly, statistically and methodologically the status of the Traditional Religious Belief, Superstition, and Witchcraft factors passed largely unchallenged. Consequently these factors retain their former identities, but I allow them to have certain correlations with the other factors as indicated below.

Lastly, given that the original evidence for subscale orthogonality in the PBS was of dubious validity, and given also that there was some evidence for moderate correlations between subscale factor items, I think it prudent statistically and conceptually to allow for correlations between factors. However, I shall initially be fairly conservative with respect to which pairs of factors I allow to correlate. Thus, I shall allow Traditional Religious Belief to correlate with Witchcraft on the grounds that for most people witchcraft is traditionally considered to be an alternative to Christian worship. Also, in accepting the existence of a devil, many people with traditional religious beliefs accept that witchcraft is one way in which the devil's power may be made manifest. It would be surprising then if there was not a correlation between the factor of religious belief and witchcraft. As I have argued, Witchcraft and Psi share a certain fundamental similarity; dependence on some unexplained or occult human ability to influence people's minds and everyday matter. Psi belief

as a folk psychological concept then may be regarded as a twentieth century expression of older magical beliefs. Consequently, I allow for a correlation between my hypothesised factor of Psychic Beliefs and Witchcraft. Finally, though superstitious beliefs are of dubious paranormality my own and Irwin's (1993) definition, some (e.g. Planer, 1988) would see superstitious beliefs as being essentially psi-like beliefs. As I remain uncertain over the status of Superstition I permitted that factor to correlate with the Psychic Belief factor but reserved judgement about its dependence with other factors, and thus constrained all other correlations with that factor to zero.

I have now outlined the substantive assumptions and reasons which lead me to postulate a radically different model of the structure of Tobacyk and Milford's PBS. Taking these considerations in to account in my specification of an actual model, I was able to produce a five factor oblique model of the PBS which I present in Figure 2[2]. Tobacyk and Milford's original seven factor orthogonal solution is represented in Figure 1. My main claim is that the Oblique Five will provide a better fit to a sample of data from the PBS than will the Big Orthogonal Seven. This claim may be tested by factor analysing both models

and comparing their fit statistics, in particular the degree to which the Oblique Five are a significant improvement over the Orthogonal Seven.

Methods

Subjects

131 new first year undergraduate students at the University of St. Andrew's filled out a large questionnaire as part of a research methods course dealing with attitude measurement. This questionnaire contained Tobacyk's (1988) Revised Paranormal Belief Scale. The students had no prior formal

---

[2] Readers should ignore the numbers placed next to arrows in Figures 1, 2, and 3 for the present. These are parameter estimates (i.e. results) and thus do not have any bearing on the specification of the model until after the results have been discussed.

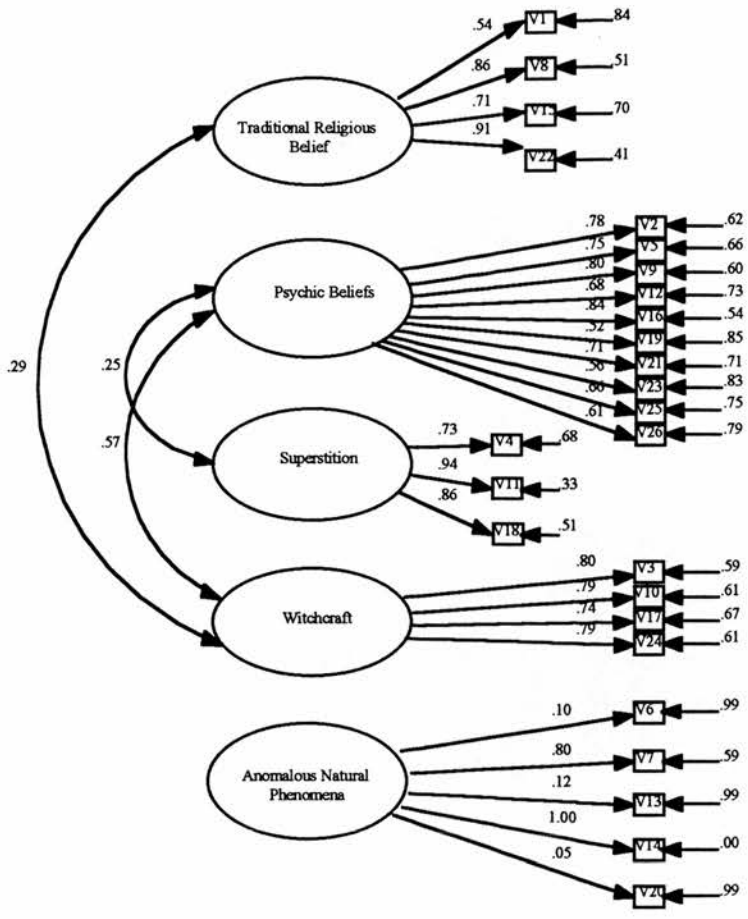

Figure 2. The Oblique Five.

acquaintance with the topic of the paranormal, or with Dr. Roe's views on the paranormal.

Materials

Tobacyk's (1988) 26 item Revised Paranormal Beliefs Scale.

Procedure and Statistical Analysis Details

A table of inter-item covariances was calculated and analysed using the technique of covariance structure analysis using EQS (Bentler, 1989). EQS combines factor analysis, multiple regression, and path analysis techniques,

providing a powerful system to test and confirm various causal or structural models, including the more basic confirmatory factor model used here.

The most important output from a test of any confirmatory factor analysis is a list of goodness of fit statistics, indicating how well the hypothesised factor model accounts for the observed correlations (or covariances) in one's data. When assessing the goodness of fit of a confirmatory factor model there are a number of statistics which we have at our disposal. For the first of these, the chi square goodness of fit statistic, we require that the chi square be *non-significant* for the model to be considered a good fit. Next we have average standardised residual statistics, of which the off diagonal residuals are the most important. These may be treated like residuals in regression analysis, indicating unexplained covariance, and are particularly useful in comparing the differences in fit between two or more alternative models. Aikaike's Information Criterion (AIC) chi-square provides an estimate of goodness of fit that takes into account the parsimony of the underlying model (in terms of the number of parameters that must be estimated to arrive at a particular solution). Thus the smaller the value the better. Lastly, Bentler and Bonett's non-normed fit index (NNFI) provides a particularly good estimate of model fit with small sample sizes (appropriate for my sample size). To even be considered as a reasonable fit, the fit index should show values of greater than 0.9. Where two competing models are to be tested one against the other, a simple test of the difference in fit between the models may be obtained by subtracting the chi-square value for the better fitting model from the chi-square value for the worse fitting model and comparing the new chi-square on *d.f.*(model 1) minus *d.f.*(model 2) degrees of freedom.

Lastly, a useful feature of EQS when developing models is the suggestion of model modifications in order to provide a better fit. Used cautiously, this feature is helpful in achieving models which fit the data well.

TABLE 1. MODEL FIT STATISTICS	Model 1: Tobacyk and Milford's (1983) Big Orthogonal Seven (see fig. 1)	Model 2: Lawrence's (1995a) Oblique Five (see fig. 2)	Model 2a: Lawrence, Roe, & Williams' Modified Oblique Five (see fig.3)
Avg. off-diagonal absolute standardised residuals	0.29	0.14	0.10
Independence Model Chi-Square (*d.f*)	2248.345 *(325)*		
Independence AIC	1598.345		
Model AIC	213.539	95.802	42.759
Chi-Square (*d.f*)	811.539 *(299)*	687.802 *(296)*	626.760 *(292)*
Chi-Square *p*-value	<0.001	<0.001	<0.001
Normal theory RLS Chi-Square	920.57	694.37	642.20
Bentler-Bonett Non-Normed Fit Index	0.710	0.776	0.806
	Significance Test of Model 1 Fit vs. Model 2 Fit		**Not Confirmatory**
Model Differences in Chi-Square (*d.f*) (*p-value*)	123.737 *(3)* *(<0.0001)*		**Not Confirmatory**

Results

Table 1 provides model fit statistics for the confirmatory test of models 1 and 2, the Orthogonal Seven and Oblique Five models respectively. The fit statistics for my exploratory modified model (model 2a) are given in the fourth column of table 1. Figures 1, 2 and 3 show the path diagrams for their respective models with standardised solution parameter estimates placed by the relevant paths of influence. The little numbered "v's" in boxes indicate the items, whilst the large named circles are the hypothesised factors. Arrows passing from the factors to the items indicate that item loads on (or is 'caused' by) that factor. The strength of this influence is given as a path coefficient (i.e. correlation) above the respective path. The open numbers connected by arrows running to the variables from the right hand side indicate the degree of error in the estimation of the variables. Squaring the values of the parameter estimates for the item loadings or unique errors indicates the percentage of variance accounted or unaccounted for (in the case of errors) by that source of variance respectively.

As we can see, Tobacyk and Milford's (1983) Big Orthogonal Seven do not provide a very good fit to the data. The average off diagonal standardised residuals are particularly poor, indicating that on average the amount of correlation between items left unaccounted for was 0.29. The non-normed fit index also falls well short of the 0.9 Figure required for the model to be regarded as even minimally acceptable. In brief, the Big Orthogonal Seven almost certainly do not represent the underlying factor structure of the PBS, even though examination of the standardised solution parameter estimates in Figure 1 show that all items load their factors adequately.

The fit statistics for model 2, the Oblique Five, are clearly a considerable improvement. The average off diagonal standardised residuals are half that of those for the Orthogonal Seven model. The non-normed fit index is higher, but still well short of being acceptable. Equally the chi-square goodness of fit statistic is highly significant ($\chi^2_{(296)}=687.802$, $p<0.001$), indicating that the model still fails to adequately reproduce the observed inter-item covariance matrix. Comparing the AIC statistics, however, for the two models we see that the Oblique Five represent a far superior fit to the data when the parsimony of the solution is taken into account. Lastly, and perhaps most importantly, the test of the difference in fit between the two models shows that the Oblique Five are a significant improvement over the Orthogonal Seven, despite the overall lack of fit for either taken separately.

Examination of the standardised solution parameter estimates show that modelling the old Spiritualism, Psi beliefs and half of the Precognition factor items as a Psychic belief factor was perfectly acceptable. In addition modelling the correlations between factors (the two headed arrows linking factors) was a good idea, and the values for the correlations are all significant according to univariate tests of parameter significance. However, despite its improvement in fit over the Orthogonal Seven, Lawrence's Oblique Five model can be improved upon. In particular, we can see that the Oblique Five provide an awful representation of the factor loadings for items 6, 13, and 20. These items

simply could not be forced to load with items 7 and 14[3] (which refer to astrological prediction). At this point I ordered a Lagrange Multiplier test on the Oblique Five solution, in order to see whether freeing up certain parameters might improve the fit of the model. Amongst the top twenty largest changes suggested four involved permitting correlations between the remaining uncorrelated factor pairs, and two of these fell in to the top ten of model modifications. Indeed the most significant improvement in model fit suggested was to free the correlation between the Anomalous Natural Phenomena factor and the Superstition factor. As I had initially been over-conservative in the modelling of factor correlations, and I felt that freeing up the correlations was intuitively plausible, I decided to implement the suggested changes by freeing up the correlations in those factors suggested.

According to MacCallum *et al.* (1992) over eager use of the Lagrange Multiplier test may result in capitalisation on chance events in the data and consequently I refused to model any further modifications. In particular three of the top ten modifications were to allow extraordinary life form items in the Anomalous Natural Phenomena scale to load on the Psychic belief, Superstition and Witchcraft factors. I chose not to implement these changes, despite the fact that significant improvements in fit would have occurred, for reasons that I outline in my discussion. Of the other five top ten modifications only the second most significant made any real sense and thus I chose to stick (some would say over cautiously) with simply freeing up factor correlations. This new modified Oblique Five model is displayed graphically in Figure 3, with fit statistics given in the fourth column of table 1.

---

[3] Item 14 is seen to load perfectly on its factor, and its error coefficient is zero. This is not in fact a genuine parameter estimate. Due to problems in identification of the Anomalous Natural Phenomena factor item 14 was constrained by EQS at its lower bound. Examination of correlations between parameter estimates revealed that values for the loadings of items 7 and 14 on the Anomalous Natural Phenomena factor had exceeded the maximum possible correlation with the estimate of error variance for item 14 (+1). These problems stem from the fact that the present model did not permit correlations between the Anomalous Natural Phenomena factor and the other factors. When these correlations were allowed in the modified model, the identification problem clears up.

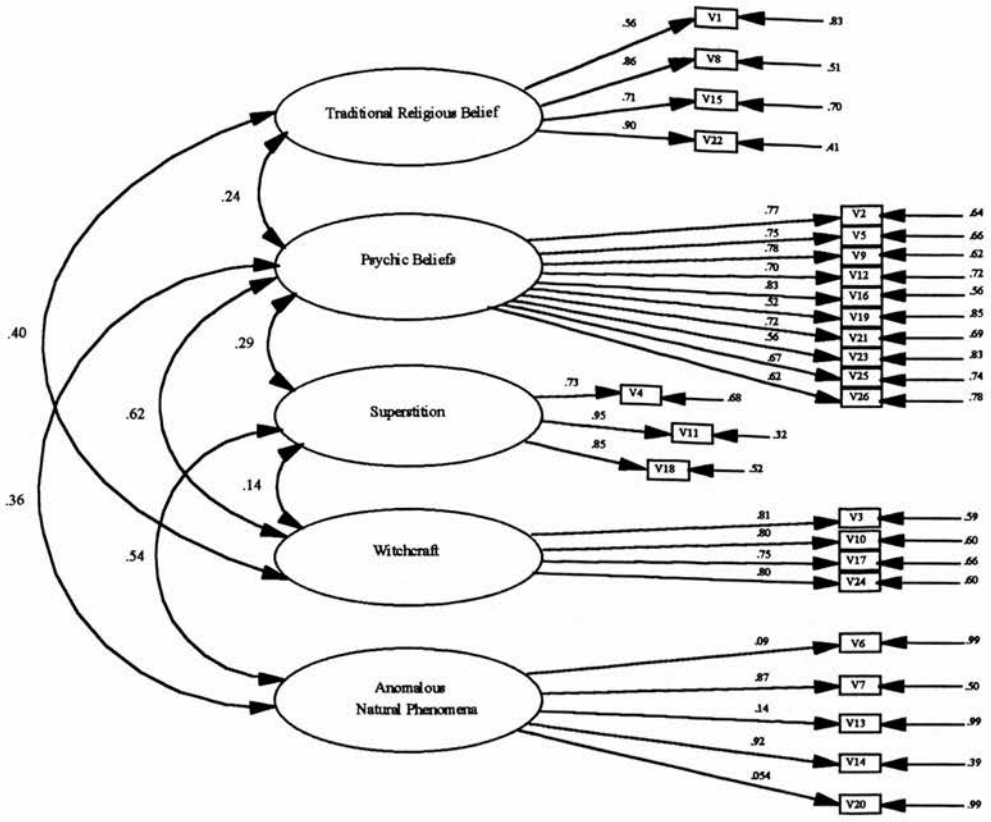

Figure 3. The Modified Oblique Five

The modified Oblique Five model obviously improves over the original Oblique Five (it was designed to do so). The improvement that results brings the model close to adequacy in terms of the non-normed fit indices, but still leaves much to be desired in respect of overall fit and residual statistics (good values for residuals are down near 0.02-0.04; see Lawrence *et al.*, 1995, for a model with excellent fit statistics). The moral of my modifications is clear, the PBS factors are most certainly oblique, not orthogonal, as Lawrence (1995a) predicted they should be.

However, I have not been able to provide an adequate model of the structure of paranormal belief. Despite these modifications a better model exists. I now turn to a discussion of the results.

Discussion - Study 1

In this study I have attempted to show that the factor structure of the PBS can be represented more accurately by a five factor oblique model, in contrast to Tobacyk and Milford's (1983) seven factor orthogonal model. My results for the test of these two models clearly show that my Oblique Five fit the data better than Tobacyk and Milford's Orthogonal Seven. Although the improvement in fit is substantial, my original Oblique Five and my modified Oblique Five still do not provide an adequate fit to the data. In short, some other model exists which provides a better representation of the structure of paranormal belief, and it remains for future research to track down that structure. Further to this I wish to suggest some ways in which one might achieve a model of paranormal beliefs which is a good fit.

In my modified Oblique Five model (MOF) it was clear that the model failed to adequately specify the factor structure and loadings for the items referring to Tobacyk's (1988) Extraordinary Lifeforms (ELF) subscale. Examination of the largest standardised residuals showed that considerable covariance between the ELF items and the Psychic Belief, Superstition, and Witchcraft items was left unexplained. A further Lagrange Multiplier test confirmed this view with three of the top five suggested modifications being the freeing up of loadings for ELF items with the above mentioned three factors. Clearly then, the problem for the MOF model is that ELF items are left dangling with nowhere to load. There are at least three solutions to this problem.

Firstly, one could simply allow the ELF items to form their own factor, thus reconstituting the factor of Extraordinary Lifeforms, which would almost certainly then correlate with Psychic Belief, Superstition, and Witchcraft factors. Secondly, one could retain the present MOF model and permit the ELF items to load on those factors suggested by my Lagrange Multiplier test. This would be equivalent to suggesting that belief in the Loch Ness monster, the Abominable Snowman, and life on other planets was caused by an underlying belief in psychic powers, superstition, and magical forces. This view

is however deeply problematic, primarily because the ELF items are intrinsically different in nature from the other factor items. Perhaps then the relationship between ELF items and the other factors is not so much causal but genuinely correlational. Thus, both psychic beliefs and ELF beliefs are beliefs about strange phenomena. The same applies to superstition and witchcraft beliefs. But if one person holds strong beliefs about, say, psychic powers (maybe on the basis of paranormal experience of oneself or close others) then that person's views about other strange (but not necessarily paranormal) phenomena become more open simply because their 'boggle threshold' has already been lowered. Effectively the paranormal believer is saying 'Well, I've had to change my views about certain strange phenomena of the paranormal sort, and I wouldn't be surprised then if the world were generally far stranger than most would think.' As a result one finds positive correlations between many diverse views about strange phenomena, of one sort or another. But this is not to imply that views about ELF are intrinsically of the same sort as views about paranormal phenomena proper.

According to this view then, permitting the ELF items to load on a mixed bag of paranormal belief factors is not appropriate. Option two above is then ruled out, even though it would improve the fit of MOF. Option 1 remains feasible in a weak sense, in that it doesn't flagrantly flout notions about the intrinsic nature of either psychic belief, superstition, or witchcraft. However, in a strong sense the continued presence of an ELF factor in a scale purporting to measure paranormal belief does present problems for the conceptual clarity of the model. Consequently, we seem to be forced to accept a last radical measure in order to improve our prospects of modelling accurately the structure of paranormal beliefs - cut the ELF items out! This suggestion is no more than the practical acknowledgement of the fact, agreed on by both Tobacyk (1995) and myself (Lawrence, 1995a) that the ELF subscale does not measure paranormal belief. Somewhat circuitously then, we are forced by the empirical data to concede to my own stronger conceptual critique of the PBS.

My final suggestion for future researchers aiming to model the structure of paranormal beliefs then, is to omit items referring to extraordinary lifeforms. Having done this for my own data in a casual exploratory way I was able to fit a reduced Oblique Four model (omitting *only* the Anomalous Natural Phenomena items) which provided an almost fully adequate reconstruction of the remaining data. However, the main confirmatory results from my research are clear - the factor structure of the PBS can be accounted for better by fewer than seven factors.

Study 2

In the last few months of my Ph.D. write up I was able to test once again[4] the Orthogonal Seven and Modified Oblique Five Models analysed, or in the case of the modified oblique five *developed*, in the study just discussed. This replication would help to confirm that the superiority of the Modified Oblique Five is not just capitalisation on chance. As the model specification process here has already been detailed in the introduction to this chapter, I shall proceed to the methods section without further ado, suffice to note that I shall be comparing the orthogonal seven with the *modified* oblique five, not the original oblique five.

Methods

Subjects

The participants were 225 students taken from four of the schools within Coventry University. These students had no prior acquaintance with Mr. De Cicco. All participants were assessed in a lecture hall as a group prior to the scheduled lecture. The age range was from 18 to 50, with the mean age being 20½ years and the median age being 19. There were 118 men and 107 women.

[4] My sincere thanks to Pio De Cicco who diligently carried out the data collection for me at Coventry University.

Materials

All participants were issued with the 26 item revised Paranormal Belief Scale (Tobacyk, 1988), and were also asked to give their sex and age.

Procedure

Once approval from the dean had been given, the author's assistant arranged to administer the PBS to an available class of students, by arrangement with the class lecturer. Four such classes were used, from four different schools within the university (Health and Social Sciences, Art and Design, Mathematical and Information Sciences, and Business Studies). All participants were given a very brief description of what was required before the PBS was handed out for completion. After completion the author's assistant thoroughly debriefed the class about the study aims, fielded questions, thanked them and their lecturer for their time, and departed.

Statistical Analysis

A table of inter-item covariances was calculated and analysed using the technique of maximum likelihood confirmatory factor analysis using EQS (Bentler, 1989).

Results

All three models tested were analysed without technical difficulties. Table 1. gives the results of the covariance structure analysis for models 1, 2 and 3, the Orthogonal Seven, Modified Oblique Five, and Oblique Four models respectively. It is clear that the pattern of results found in Lawrence et al. (1997) confirmatory factor analysis of the PBS are repeated here. Thus, the Orthogonal Seven Model advocated by Tobacyk and Milford (1983) gives an abnormally large average off diagonal residual. This indicates that the Orthogonal Seven fail to explain much of the correlation between items in our sample data. The Modified Oblique Five leave, on average, half as much to be explained in comparison to the Orthogonal Seven. The chi-square goodness of fit test indicates that both models 1 and 2 fail to accurately reproduce the

TABLE 2. MODEL FIT STATISTICS	Model 1: Tobacyk and Milford's (1983) Big Orthogonal Seven	Model 2: Lawrence, Roe, & Williams' Modified Oblique Five	Model 3: The Oblique Four
Avg. off-diagonal absolute standardised residuals	0.21	0.10	0.08
Independence Model Chi-Square (d.f)	2832.313 *(325)*		2260.830 *(210)*
Independence AIC	2182.313		1840.830
Model AIC	601.021	284.320	186.349
Chi-Square (d.f)	1199.021 *(299)*	868.320 *(292)*	554.350 *(184)*
Chi-Square p-value	<0.001	<0.001	<0.001
Normal theory RLS Chi-Square	1454.629	861.303	562.468
Bentler-Bonett Non-Normed Fit Index	0.610	0.744	0.819
	Significance test Model 1 vs. Model 2 Fit		
$\Delta\chi^2$ *(df) (p-value)*	330.701 *(7) (<0.001)*		

sample correlational data, but the Modified Oblique Five are markedly less discrepant in this regard, and the chi-square difference test ($\Delta\chi^2$) indicates that the Modified Oblique Five are significantly more accurate in their reconstruction of the sample correlational data. The AIC chi-square tests indicate that the modified oblique five represent a more parsimonious depiction of our sample data. Finally, the Bentler-Bonett non-normed fit index

shows that the modified oblique five improve in their level of explanation over the orthogonal seven, but as neither of these models reach the 0.90 minimum for this index we cannot assume that an acceptable representation of the PBS has been found. The results for the Oblique Four model show that when extraordinary lifeforms and astrology items are excluded, the remaining specification of items on their factors is less discrepant still. Comparatively, the average residual, AIC and non-normed fit indices for the Oblique Four show that this model provides the best reconstruction of the data *relative to its dataset* than does either the orthogonal seven or the modified oblique five relative to

theirs. Nevertheless, the NNFI still indicates an unacceptable level of incremental model fit.

Discussion - Study 2

The results of this study confirm Lawrence's (1995a) critically inspired suggestion and Lawrence, Roe and Williams' (1997) factor analytic evidence for a five factor oblique structure to the PBS. This confirmation is a relative statement of superiority however, as the fit statistics for all models tested in this study and in Lawrence et al. (1997) study fall short of acceptance. Thus, the PBS remains resistant to any straightforward representation of its factor structure, even when construct validity 'challenged' items about extraordinary lifeforms and Abominable Snowmen are removed. Given these problems, one feels inclined to emphasise the need for a wholly new measure of paranormal belief. For whoever chooses to develop it, a word of advice; go obliquely and remember, seven is not the magic number.

Conclusions

My initial interest in the PBS began as a result of a need to pick a measure of paranormal belief for use in my own research. I perused the various options and realised that of all of these the PBS was the *least* problematic. However, on deepening my analysis of the PBS I came to realise, as outlined in chapters four and five, that even the PBS was not satisfactory. My initial critique of the PBS was largely based on the application of certain basic principles in psychometrics, and as such had precious little to do with any empirical assurance that something was wrong with the scale. Thus, a truly convincing demonstration of my case could only pertain if new evidence could be brought to bear upon the number of factors problem, and the problem of orthogonality in the PBS. In my initial critique of the PBS I had laid my cards on the table, as it were, by suggesting that a more appropriate factor solution in the PBS might have given fewer than seven factors, and that these would most

certainly by correlated (i.e. non-orthogonal). In keeping with my advocacy of CSM to resolve problematic areas of theory in this area, I chose to perform a confirmatory factor analysis in order to compare empirically the validity of Tobacyk's Big Seven and my own preferred 'flavour' the Oblique Five. CFA is, after all, just CSM without the structural equation model component. Using this tool, we can see that my expectations were fulfilled. This confirmation does not show that my Oblique Five are correct, only that they are less inadequate than Tobacyk's Big Seven.

Chapter 6

PRELIMINARY ATTEMPTS AT MODELLING: DEVELOPING THE CHILDHOOD FACTORS MODEL.[1]

"When it's grey I know all I want to do - resort to make believe."

(From the song A Love for all Seasons, by Right Said Fred)

Any particular person's history of paranormal belief and experience must have some beginning, and some causative factor responsible for that beginning. Until quite recently researchers on the topic of paranormal belief (PB) and experience (PE) remained unaware, or uninterested, in the possible early developmental causes of PB and PE, much of the research in this area reflecting, as we saw in the introduction, what Irwin (1993) calls the *cognitive deficits hypothesis*- the view, influenced by a broadly sceptical research agenda, that those with high PB or PE are '...illogical, irrational, credulous, uncritical, and foolish' (Irwin, 1993, pp.16). Recently, however, research conducted by Irwin (1992) has shown the potential importance of childhood factors underlying PB, and so it is to a subset of these potentially causative correlates of PB and PE that this chapter is addressed- namely, childhood trauma and childhood fantasy.

Research addressing the childhood causes of PB and PE can be traced back to a small study reported by Wilson and Barber (1983). In their study Wilson and Barber (1983) interviewed 27 women whom they had rated as excellent hypnotic subjects (with a comparison group of 25 women not so rated). The interviews centred around childhood and adult memories, fantasies and psychic

---

[1] I wish to thank Claire Edwards, Nicholas Barraclough, Sarah Church and Francesca Hetherington for diligently carrying out the research reported herein as part of their second year project (supervised by myself). All the analyses and interpretations are mine however. This chapter may be found in published form as Lawrence *et al.* (1995) in the journal *Personality and Individual Differences,* Vol. 19, No. 2, pp. 209-215.

experiences. As Wilson and Barber (1983) note, with reference to their excellent hypnotic subjects,

> *"Their extensive and deep involvement in fantasy seems to be their basic characteristic and their other major talents—their ability to hallucinate voluntarily, their superb hypnotic performances, their vivid memories of their life experiences, and their talents as psychics or sensitives—seem to derive from or to grow out of their profound fantasy life." (Wilson and Barber, 1983, pp.340).*

One of the most significant factors in differentiating between fantasisers and non-fantasisers in Wilson and Barber's group was the incidence of childhood trauma (previously found by Hilgard (1970, 1974)). Of the 27 fantasisers, nine (33%) reported having experienced some difficult or stressful event in their early life, ranging from physical abuse by parents, or a mother with emotional problems, to unstable living conditions. No subjects in the comparison group reported having experienced any childhood trauma (see Lynn and Rhue, 1988, for a replication).

Since Wilson and Barber's study, two studies relating fantasy proneness to PB have been conducted by Irwin (1990; 1991b). In the first study, Irwin (1990) administered Myers' (1983) version of the Inventory of Childhood Memories and Imaginings (ICMI:C) to 92 people, along with Tobacyk and Milford's (1983) Revised Paranormal Belief Scale (PBS). Irwin found Pearson correlations of .48, .58, and .43 for the Psi Belief, Spiritualism, and Precognition subscales of the PBS respectively with scores on the ICMI:C (all significant at $p<0.001$), suggesting a definite link between imaginal involvement and the core constructs of PB. In a follow up study, Irwin (1991b) administered the ICMI:C and PBS to a sample of 122 students and found a similar pattern of correlations- the PBS Psi Belief, Spiritualism, and Precognition subscales

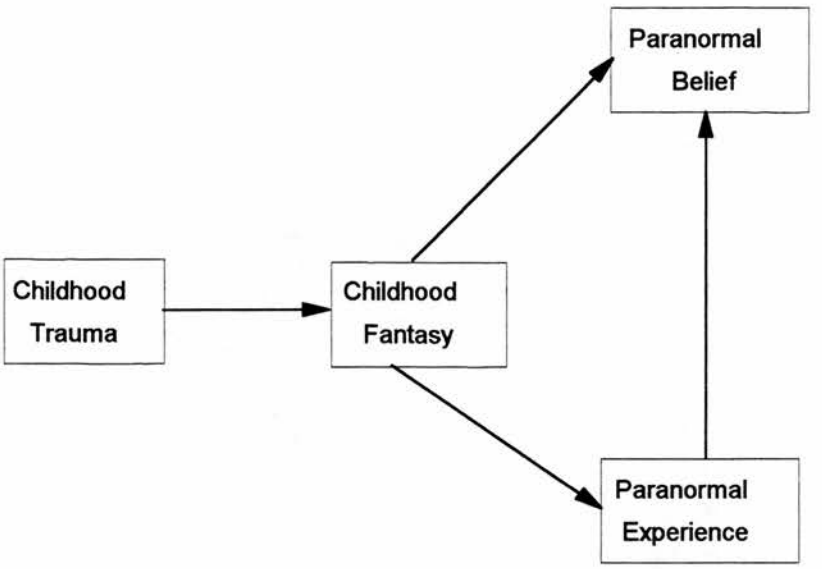

Figure 1. A simplification of Irwin's (1992) model depicting the childhood causes of paranormal belief.

correlated .40, .42, and .34 with the ICMI:C (all $p<0.001$). Another study by Rao (1992) looked at the relationship between fantasy proneness and PE alone as measured by Neppe's (1983) Subjective Paranormal Experiences Questionnaire (SPEQ). Rao found a correlation of .46 ($p<0.01$) between the SPEQ and the ICMI:C, replicating Wilson and Barber's finding that fantasy prone participants reported more psychic experiences.

On the topic of childhood trauma and PB Irwin (1992) administered the revised PBS (Tobacyk, 1988), and the Survey of Traumatic Childhood Experiences (STCE) (Council and Edwards, 1987) to 72 students. He found a Pearson correlation of .21 ($.05<p<.08$) between the full scale PBS score (including items of dubious paranormality about the Loch Ness monster and life on other planets) and the STCE; a non-significant result which is at least suggestive. Those items most strongly correlated with PB were the STCE items relating to intrafamilial physical abuse ($r= .31$, $p<.05$). Thus there is evidence to support the view that childhood trauma has an influence on the subsequent

development of PB. In suggesting his own model of these relationships Irwin concludes,

> *"...Fantasy proneness...is a central mediating variable in the development of paranormal beliefs. Fantasy proneness may be kindled in early childhood through parental encouragement and/or the incidence of traumatic experiences. In particular, the latter will instil a strong need for interpersonal control. Given a conducive social context, fantasy proneness and the associated need for control may lead to the formation of paranormal beliefs. Fantasy proneness...also may be a factor in the occurrence and/or acknowledgement of parapsychological experiences." (pp. 206)*

Figure 1 presents a reduced version of Irwin's *childhood factors model*, showing only those relations between factors that I have chosen to model in the present study. With respect to his larger model Irwin notes, "Most of the relationships indicated by the arrows have now been documented in the empirical literature, but there remains a need for a single study entailing a path analysis of all variables encompassed by the model." (Irwin, 1992, pp. 205).

In this study I attempt a partial confirmation of Irwin's model, using a more advanced technique than path analysis, namely covariance structure modelling (CSM) using EQS (Bentler, 1989), a program specifically designed to implement and test structural equation models using covariances. Using the reduced version of Irwin's model (see fig. 1) I aim to test the causal structure underlying the relations between childhood trauma, childhood fantasy, PE and PB. No study has to date looked specifically and exclusively at the role of *childhood* fantasy involvement in the development of PB and PE, but I feel that an emphasis on childhood fantasy, as opposed to the mixed effects of child and adult fantasy (as measured by the ICMI) is important given the proximity of childhood fantasy, and its pertinence to traumatic events in childhood.

Method

Subjects

The respondents were 82 students and friends of the author's collaborators, all at the University of Edinburgh. Unfortunately no information on sex or age was taken, but the sample was broadly similar to those that one usually gets from university students.

Materials

A 61 item questionnaire was constructed using items from the following three scales;

(1) Survey of Traumatic Childhood Experiences (STCE) -- This scale, developed by Council and Edwards (1987), comprises 30 items dealing with events that may cause trauma in a child's life. The full 30 item scale was used in this study. The scale has items relating to intrafamilial sexual and physical abuse, loss related to friends and family, extrafamilial sexual abuse, assault, loss of home and property, isolation and personal malady, adolescent abortion/miscarriage, and parental divorce/separation. Responses are scored on a 5 point scale of frequency, with (A) none, (B) one, (C) two to five, (D) six to ten, and (E) more than ten traumatic experiences.

(2) Australian Sheep-Goat Scale (ASGS) -- This scale comprises 18 items referring to three common components of PB; ESP, PK, and Life after Death. Thalbourne and Delin (1993) report reasonable reliability and some evidence showing concurrent validity. Items refer to PB and PE and were marked by a 7 point scale.

(3) Childhood fantasy—This construct was measured using only those items from Myers' (1983) ICMI:C that refer explicitly to childhood fantasy. Items are scaled dichotomously, in the true/false format. High scores indicate high fantasy.

Procedure

Participants were given the three-scale questionnaire to fill in, along with a blank envelope for the completed questionnaire to be placed in to preserve the participant's anonymity. Participants were asked to fill in the questionnaire as honestly as possible. Once the questionnaire was completed and sealed in the envelope the participant was briefed about the nature of the study. No participants reported having any particular problem with the content of the scales, though same pointed out that the Australian Sheep-Goat Scale was quite contorted in its wording. Once the data had been collected they were entered in to Statview for calculation of the correlation matrix, before being entered into EQS for covariance structure modelling. At this stage two participants' questionnaires were excluded from the analysis, as they constituted extreme outliers on Trauma, and one of the subjects had (amongst other unlikelies) claimed to have had more than ten abortions/miscarriages! Consequently, I suspected that at least one of the two outliers was not responding honestly. The remaining sample comprises 80 individuals.

Statistical Methods

Correlational data from this study were analysed using the technique of covariance structure analysis using EQS (Bentler, 1989). EQS combines factor analysis, multiple regression, and path analysis techniques, providing a powerful system to test and confirm various causal or structural models. The most important output from a test of any causal model is a list of goodness of fit statistics, indicating how well the hypothesised model accounts for the observed correlational structure of the data. When assessing the goodness of fit of a causal model there are a number of statistics which we have at our disposal. For the first of these, the chi square goodness of fit statistic, we require that the chi square be *non-significant* for the model to be considered a good fit. Next we have average standardised residual statistics, of which the off diagonal residuals are the most important. These may be treated like residuals in regression analysis, indicating unexplained variance, and are particularly

Table 1. Pearson correlations between the of the main variables

Variable	Trauma	Fantasy	Experience
Fantasy	.356***	—	—
Experience	.343***	.315***	—
Belief	.215*	.271*	.534*****

* p<0.06>0.05, ** p<0.05, *** p<0.01, *****p<0.0001

(Standard Deviations: Trauma = 8.13, Fantasy = 6.46, PB= 5.75, PE= 9.55)

useful in comparing the differences in fit between two or more alternative models. Aikaike's Information Criterion (AIC) chi-square provides an estimate of goodness of fit that takes into account the parsimony of the underlying model (in terms of the number of parameters that must be estimated to arrive at a particular solution). Thus the smaller the value the better. Lastly, Bentler and Bonett's non-normed fit index (NNFI) provides a particularly good estimate of model fit with small sample sizes (appropriate for this sample size). To even be considered as a reasonable fit, the fit index should show values of greater than 0.9.

Lastly, a useful feature of causal modelling is the suggestion of model modifications, whereby EQS suggests changes to the model in order to provide a better fit. Used cautiously, this feature is helpful in achieving models which fit the data well.

Results

From tables 1 and 2 we can determine how well the study has replicated previous findings on the relation between PE and belief, with fantasy and trauma. In table 1 we may note that the correlation between fantasy and PB is r=.27 (p<.05). This fits reasonably with the studies of Irwin (1990; 1991b) showing average correlations of around .44 between the ICMI:C and the PBS. The relationship between childhood fantasy and PE was slightly stronger, r=.31 (p<.01). This result offers a fair degree of agreement with the findings of Rao (1992) with the ICMI:C (r=.46). Further, these correlations for childhood fantasy are likely to be smaller simply because we were solely measuring

childhood fantasy (the ICMI:C actually has adult fantasy questions too). Obviously one cannot expect memory based estimates of childhood fantasy to correlate as highly with present day PB or PE, as would contemporary fantasy estimates.

With respect to childhood trauma and PB the correlation exactly replicates that of Irwin (1992). This is particularly encouraging, given the use of a much narrower measure of PB. However, we failed to replicate Irwin's *post hoc* finding that the strongest specific traumatic correlate of PB was intra-familial physical abuse. The data, presented in table 2, show that the strongest predictor of PB were the items relating to 'loss of property' (scores for this predictor stemmed mostly from answers to an item relating to the frequency of moving home as a child).

Table 2 gives the descriptive statistics for childhood trauma, split into its various sources. Whilst all of the four main variables gave reasonable approximations to a normal distribution, this was not the case for the specific traumatic variables presented in table 2. Of these, only intra-familial physical abuse data showed evidence of being normally distributed, the other variables being more chi-squared in their distribution. Indeed, it is clear that most participants had neither experienced much childhood trauma, nor collectively displayed much variation in trauma experienced. This is particularly the case for familial sexual abuse, almost all participants reporting no sexual abuse in their family as children. Consequently, for correlational analyses on the four main variables I used Pearson product moment correlation coefficients, but for analyses on specific traumatic factors I chose to use Spearman's rho, corrected for ties.

Lastly, the correlations between the various factors of childhood trauma and PB and experience are interesting in so far as they show that the strength and significance of the correlations reported are much stronger between trauma

Source of Childhood Trauma	Mean	S.D.	Corrected Spearman rho (Belief)	Corrected Spearman rho (Experience)
Intra-familial sexual abuse	1.04	0.16	-.008	.058
Intra-familial physical abuse	2.49	0.92	.196 *	.194 *
Loss (friend)	1.28	0.45	-.024	.23 **
Extra-familial sexual abuse	1.28	0.47	-.038	.086
Loss (family)	1.63	0.56	.110	.164
Isolation	1.35	0.67	.064	.056
Personal Illness	1.40	0.54	-.039	.194 *
Parental Divorce/Miscarriage	1.09	0.19	.031	.137
Assault	1.57	0.69	-.094	.18
Loss of Property (Home Moving)	1.52	0.40	.379 ****	.239 **
Robbery/Mugging	1.12	0.37	.065	.251 **

* $p<0.09>0.05$, ** $p<0.05$, **** $p<0.001$

Table 2. Descriptive statistics for specific sources of childhood trauma, with correlations given for paranormal belief and paranormal experience.

and experience than for belief. Here we see the inklings of a model relating these factors, and it is to the modelling of these factors proper that I now turn.

Results of Causal Modelling

First I ran EQS on a slightly simplified version of Irwin's (1992) model, given in figure 1. The results for this model, in terms of fit statistics, are given in the appropriate column of table 3. The chi-square of 5.544 on 2 degrees of freedom indicates a near significant lack of fit between the model and the observed data. Equally the non-normed fit index is poor; all the worse given that this index is best for reflecting fit at low sample sizes. In short, Irwin's model is a poor fit.

EQS recommended two changes to Irwin's model. Firstly, the Wald test for dropping parts of the model suggested the exclusion of the direct causal link from fantasy to belief. Secondly, the Lagrange Multiplier test for adding parts to the model suggested the inclusion of a direct causal link between trauma and PE. Consequently, I made these changes to the model, arriving at a much more

Table 3. Fit Statistics for Irwin's (1992) childhood factors model and the new model

MODEL FIT STATISTICS	Irwin's Model (fig.1)	New Model (fig.2)
Average off-diagonal absolute standardised residuals	0.058	0.022
Independence Model Chi-Square (df)	52.332 (6)	
Independence AIC	40.332	40.332
Model AIC	1.544	-2.691
Chi-Square (df)	5.544 (2)	1.309 (2)
Chi-Square p-value	0.0625	0.520
Normal theory RLS Chi-Square	5.354	1.298
Bentler-Bonett Non-Normed Fit Index	0.771	1.045

streamlined model whereby trauma causes fantasy causes experience causes belief, with a direct link from trauma to experience (I had actually considered modelling this model previously, minus the trauma-experience link).

The results for the test of this model are given in the last column of table 3. The new model is a considerable improvement. Residuals and fit statistics indicate an excellent degree of fit. AIC statistics, plus residuals, show that this model is considerably better than Irwin's in reproducing the covariance found in the original data. In figure 2 I provide a graphical representation of this new model, with path coefficients. It is to a discussion of these results that I now turn.

Discussion

The correlation between childhood trauma and PB, whilst of marginal significance, was exactly the same as that found by Irwin (1992). Also the previous findings of a relation between fantasy, PB and PE were confirmed, whilst also being extended to the influence of childhood fantasy.

More importantly, Irwin's proposed model of the causes of PB, as it applies to the variables I studied, did not provide a good fit to the data. The data provided insufficient support for the link between fantasy and PB, and it certainly makes more sense to suppose that any direct effects of heightened

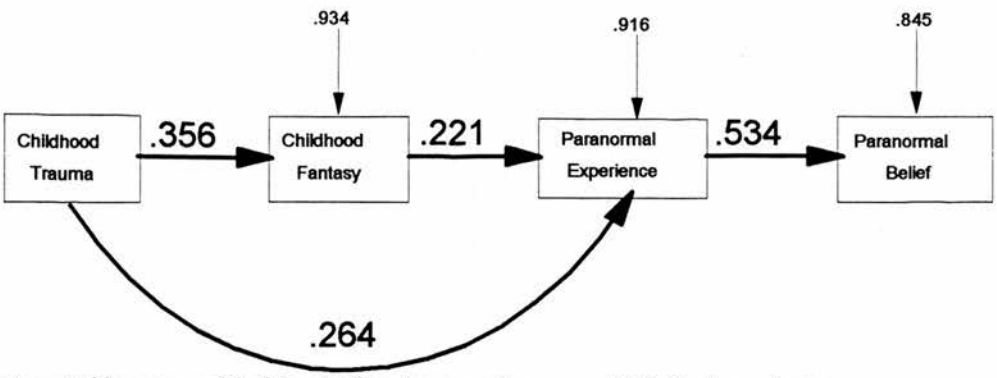

Figure 2. The new model of the childhood causes of paranormal belief and experience.

fantasy involvement would be felt upon PE, rather than PB. however, of particular interest is the finding that a much better fit stems from the inclusion of a direct link between trauma and PE, not an obvious result by any means. According to the new model trauma has two causal routes in its influence on PE, one direct, the other indirect.

The finding that trauma gives rise to PE directly, is something which one may explain only once one has been able to determine the specific traumatic causes responsible for the direct route to PE. From a *post hoc* perusal of the data I would like to suggest a possible cause of the two routes to paranormal experience (from trauma). It may be that the direct link stems from the victims' response to bereavement related trauma. A pervasive wish to bring back a loved one may be supposed to increase the likelihood of hallucinating their presence (indeed, studies of bereavement have shown this to be the case for some people, e.g. Rees (1971)). The indirect route to PE via childhood fantasy almost certainly stems from the victims' concerns about the loss of control of social and domestic domains, as Irwin's larger model suggests. Physical and sexual abuse, frequently moving home, long periods of isolation due to travel or illness, may all leave a child feeling as though they are powerless to control events in their life. In the struggle to regain control, some may turn inwards to their own uniquely private inner life of fantasy friends and make believe realms,

and in doing so they may permanently open doors to heightened visualisation ability.

Also, the new model shows that PB is best modelled as a consequence, rather than a cause, of PE. This is an important result, if it can be substantiated by further research on larger samples, precisely because Irwin's original model is equivocal with respect to the effects of fantasy on PB and PE, and PB and PE themselves are supposed to show reciprocal causation with one another. If, as the model suggests, PE experience is more the result of fantasy proneness, and PB is more the result of PE, then this suggests a more phenomenological explanation for PB, over a more attributional one (i.e. that one has PB and then attributes paranormality to hallucinatory experiences). In short, where PE has previously been considered the younger brother to PB, my research suggests that it is more the 'paternal cause' of PB. Whether these findings hold true is something only further studies could show. The next two chapters describe additional work seeking to confirm and extend this initial study.

Chapter 7

CONFIRMING CHILDHOOD CAUSES OF PARANORMAL BELIEF AND EXPERIENCE: A SMALL BUT SUCCESSFUL REPLICATION

"I felt it existed solely for me to find strength, in order to survive the world of reality...it was my outlet from the hells and pains of the real world."

(Deborah, traumatised as a child, on her fantasy world).

In the previous chapter I reported the results of my first attempt to examine the role that childhood trauma and fantasy play in the development of paranormal experience and belief (Wilson and Barber, 1983; Irwin, 1992). As I hope is now clear, research in this particular area represents a significant theoretical departure from standard research on the psychology of paranormal belief and experience, focusing as it does on the *psychodynamic* functions that paranormal beliefs and experiences may serve for the individual. Thus, whereas much present research asks the question 'How is it that people can believe in the paranormal?' and goes on to look at how differences in believers now might explain their beliefs, this new area of research asks the question 'What role might paranormal experiences and beliefs serve in the life of the individual?' and goes on to look more closely at various factors in a person's life that might be responsible for their present day beliefs or experiences.

As we saw, according to Irwin, a key role in the development of paranormal experience and belief is played by childhood fantasy. A child's level of fantasy may be triggered by a number of things; innate ability, parental encouragement of fantasy (or fantasy promoting activities such as art), and also by incidence of childhood trauma. But the upshot of all this is that fantasy plays a causal role in the development of paranormal belief and experience. For Irwin then fantasy is causally equivalent in its effects on both paranormal belief and experience. It

can cause both with equal measure. The role of fantasy is further enhanced in Irwin's model by the fact that it is the 'central mediating' variable for all childhood influences on the subsequent belief and experience. The effects of trauma in childhood upon subsequent paranormal experience and belief, in Irwin's model, are only felt indirectly through its influence on fantasy. A similar pattern of causation is postulated once again for Irwin's other main childhood cause - parental encouragement of fantasy (though here it must be said the association is obvious). Though Irwin's full model encompasses a wide range of variables, on a variety of levels; both social and cognitive, the model tested here is simply a modified version of Irwin's original model that nonetheless expresses its essence in the central role for fantasy. This model is presented, using the typical format for presenting path models, in Figure 1. Boxes represent measured variables, and circles represent latent or factored variables. Arrows indicate the implied direction of 'causal' effects.

In the last chapter we found that Irwin's model provided a poor degree of fit to the data, in particular the relationships between paranormal experience and trauma, and paranormal belief and fantasy were misspecified. Post-hoc exploratory specification searches lead to a modified model in which a direct cause of childhood trauma on paranormal experience was the most novel suggestion. By implementing the suggested changes a near perfect degree of fit was obtained. According to this new model childhood trauma has two routes to the development of paranormal experience. The first route is via the mediation of childhood fantasy, the route that Irwin first explicated in any detail. The second route, however, is a genuinely new development. Assuming there to be no mis-specification present in the new model, then childhood trauma also has a *direct* effect upon the development of paranormal experience. In short, simply being traumatised has effects upon subsequent development of paranormal experience devoid of any other mediating influence. If I hit you (as a child) you are much more likely to develop paranormal experiences. If this is a genuine finding then it should recur in future studies testing the new model.

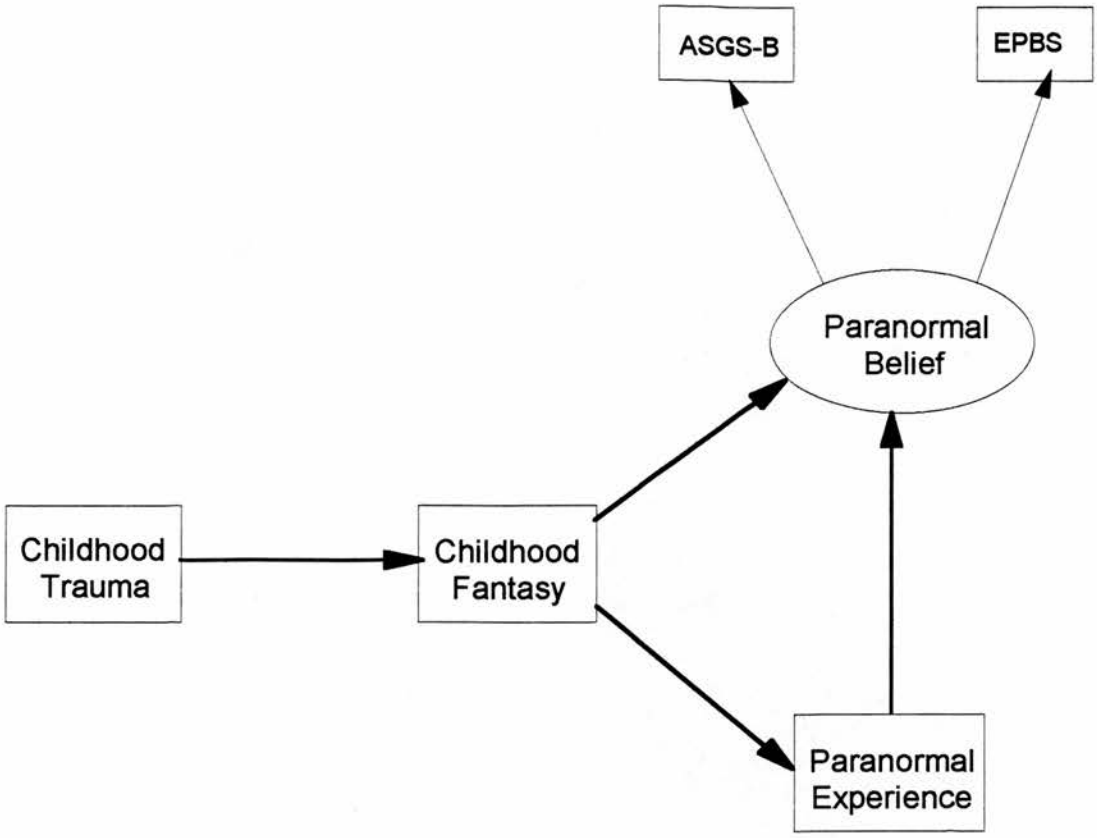

Figure 1. Irwin's (1992) model of paranormal belief development.

Of course, there is one possible problem with the previous study that threatens the validity of its findings; the new model was based upon a specification search. EQS simply searched exhaustively through all possible unspecified model parameters for any additional parameters that could be added in order to improve model fit. The link between trauma and paranormal experience was just such a parameter, plucked blind to its conceptual import on the basis of purely empirical factors. The problem with such a procedure is that it just as easily selects spurious significant parameters as meaningful/genuine significant parameters. The link between trauma and experience may therefore represent nothing more than a type I error.

If the link between trauma and experience is spurious, the result of mere sampling error, then this particular link will fail to be significant in a confirmatory test of the model (and the new model will likewise suffer some decrement in explanatory power). In this chapter then, I outline the results of a second study which served the sole purpose of enabling a confirmatory test of the new model's explanatory superiority over that of Irwin's model. An additional feature of Irwin's wider model is his suggestion that a need for interpersonal control mediates the link between trauma and paranormal belief. Irwin operationally tested this hypothesis by correlating Schultz's FIRO-B expressed control measure with paranormal belief. He found that expressed control on the FIRO scale correlated positively with paranormal belief only when the repondent's friend completed the scale. This correlation was not evident when the respondent completed the control measure. In this study the full FIRO-B scale was administered as part of the questionnaire package, and we shall therefore be able to provide evidence confirming this link (or not), as well as looking at the influence of the other five components of the FIRO-B scale. In addition, this new study served to help me pilot the larger questionnaire that is to form the basis of my final study. The present study was not without its problems however. It was to initially serve as a comparison of the beliefs of sceptics and believers. However, I was unable to enlist the support of the 100 sceptics I needed for the study, and thus was left with the responses from my sample of the SPR membership.

Method

Subjects

The respondents were 58 members of the Society for Psychical Research, part of an initial random sample of 100 members. Ages ranged from a minimum of 17 years to a maximum of 87 years. Average age was 46.03 years. Of the 57 respondents who gave their sex, 42 were men (74%) and 15 were women

(26%). The SPR membership is well known for its predominance of male members in the middle to upper classes, and the list of sample occupations reflects this (e.g. a random selection of occupations taken from the sample are labourer, university lecturer, animal carer, part time nurse/tutor, and teacher). As for the sample's representativeness of the range of paranormal belief and experience, its membership may broadly be classed as non-psychic believers in the paranormal. Therefore the nature of the sample warrants some caution when it comes to interpretation of results.

Materials

All survey participants received the 225 item survey questionnaire, and a single copy of the 54 item FIRO-B questionnaire. The nature of these questionnaires is as follows;

1. Survey questionnaire (see Appendix II) - this is a 225 item questionnaire using a variety of response formats, comprising the following scales;

> a) Paranormal Experiences Scale (PES) (see Appendix I) - this scale was put together by the author to comprehensively measure a person's experience of parapsychological phenomena. Telepathy, clairvoyance, precognition, OBE's, mediumship, and psychokinesis are all assessed.
>
> b) Modified Australian Sheep-Goat Scale (ASGS-B) - this scale was a shortened version of Thalbourne and Delin's (1994) ASGS. Only those items referring to paranormal belief were used (the ASGS also refers to paranormal experience). Items were drastically abbreviated to make their meaning clearer on the advice of colleagues whose own participants voiced some confusion about the nature of the original wording for items in the ASGS.

c) Edinburgh Paranormal Belief Scale (EPBS) (see Appendix I) - this scale was drawn together by the author in order to overcome some of the problems with other scales for the measurement of paranormal belief. The intention was not to produce a psychometrically impeccable instrument, but merely to have a second tool for the assessment of paranormal belief that might complement the ASGS (thus cancelling out their specific deficits).

d) Survey of Traumatic Childhood Experiences (STCE) - This scale remained the same as that described in chapter 6.

e) Childhood fantasy (see Appendix I) - This scale remained the same as that described in chapter 6.

2. FIRO-B - this 54 item scale was developed by Schultz (1954) to measure the fundamental determinants of interpersonal behaviour as Schultz saw them; affect, control, and inclusion. Each of these three aspects of interpersonal behaviour is measured in terms of both wanted and expressed behaviour. Thus, the wanted affect scale measures the degree to which a person wants others to act 'warm and friendly' to them, whereas the expressed affection scale measures the degree to which the respondent expresses 'warmth and friendliness' in their own behaviour. For the purposes of this study, the most interesting component of the scale is the control dimension. This scale measures the extent to which people seek to control, or wish to be controlled by, others. Respondents answer each item on a 6 point scale (with varying meaning). These raw item scores are then filtered through a binary scoring system in which there are 0-9 points available for each of the six measured constructs (wanted/expressed affection, wanted/expressed control, and wanted/expressed inclusion). Schultz has identified 37 characteristic scoring patterns, which he has wryly named, e.g. Loner, Mission Impossible. Each

pattern identifies a characteristic way of dealing with that aspect of interpersonal behaviour. Furnham (1990) describes the scale as having, 'limited, but satisfactory psychometric properties in terms of validity and reliability.' (pp.712). The primary area of use for the FIRO has been in the area of occupational selection.

Procedure

A random sample of 100 members of the SPR was drawn up from a list of over 700 British members in the 1994 membership lists (kindly provided by Eleanor O'Keeffe, secretary at the SPR). Copies of the survey questionnaire were then posted to the sample over the period 8 August 1994 until 22 September 1994. The retrieval procedure was to allow one month for completion and return of the questionnaire before a follow up phone call was made. 43 responses were received in the first posting, and 15 other responses were received after the follow up phone call. Of the original 100 people sampled, four were physically unable to respond (due to death, not living at the listed address, or some other similarly irretrievable circumstance). The final plausible response rate was therefore 60.4%. In addition to assuring participant anonymity, the author gave all potential respondents the opportunity to contact him directly by phone if they had any problems with the content of the childhood trauma scale. No phone call or complaint about the content of the scale was ever received, although several respondents expressed bewilderment at the association of trauma with paranormal experience and belief (all such queries were personally answered by letter).

Statistical Methods

The statistical analysis adopted for this study was much the same as that for the pilot study. Two differences are worth noting. Firstly, in this study two measures of paranormal belief are used, the ASGS and the EPBS. The variable

of paranormal belief in this study is therefore analysed as a latent variable defined by factor scores based on the systematic covariation in the two manifest variables. Secondly, given that non-normality is more likely in smaller samples, I chose to employ the maximum likelihood estimation technique (as before), but with robust statistical estimators selected. These estimators are adjusted to protect against the problems of non-normality that may befall studies with small sample sizes.

Results

Prior to analysis data were inspected for missing values both by eye and by graphical techniques provided by EQS. Values that appeared to be missing at random were replaced by mean substitution. The only data which appeared to be missing non-randomly were the FIRO-B data; due to its particularly repetitive response format some individuals simply stopped filling it in, or didn't bother. In analyses with the FIRO: B only complete cases have been analysed. However, with a small amount of mean substitution it was possible to recover complete data for all the main variables of interest in this study.

Firstly, examining table 1 we can see how well the basic relationships uncovered in the last study have been reproduced in this study. Of course, we can also see how well this study replicates the work of Irwin (1992). Firstly, the correlation between the two belief measures and fantasy (average Pearson r = 0.225) accords very well with the results found in the previous study based on a student population (r = 0.27). Both these correlations and those of the previous study are much lower than those found by some others using the ICMI:C; Rao (1992) found a correlation of 0.46 between the ICMI:C and paranormal experience. As argued before (Lawrence et al., 1995) one can expect the present correlations with fantasy to be below those found in other studies simply because the measure of fantasy used here addresses only childhood fantasy, and not child and adult fantasy (as does the full scale

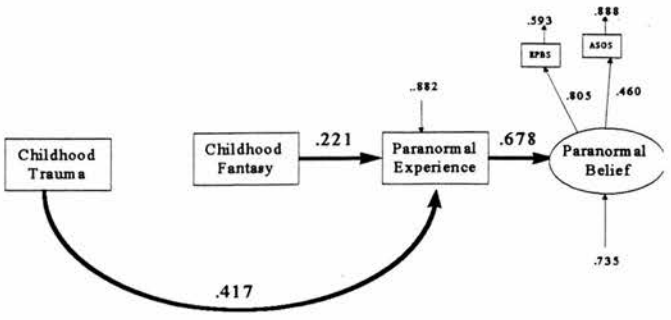

Figure 2. Lawrence et al. (1995) model of paranormal belief development.

ICMI:C). Thus, the results for the relationship between fantasy and belief show the kind of chronological attenuation that one would expect for temporally distant influences. As before, and in a similar vein, the correlation between fantasy and experience accords well with that of the previous study (r = 0.315). The consistency in these correlations between two quite distinct populations is heartening, and tends to suggest that the differences that do exist between the Edinburgh student study and the SPR study have no systematic influence upon the relationships between the variables of concern to me in this thesis.

Turning to childhood trauma and belief the degree of consistency with previous studies is lower. Whilst the EPBS does show a similar strength of relationship with trauma with Irwin's study and the previous study (r = 0.215 for both those studies), the ASGS-B shows almost no correlation with trauma. This is perhaps not so troublesome though, given that previous modelling attempts suggested that there was no significant direct effect of trauma on paranormal belief, and what is more important for modelling is that experience shows a stronger relation to trauma than does belief. This latter correlation is clearly consistent with previous results, the figure of 0.440 being somewhat larger than that found in the previous study (r = 0.343). Summarising the conclusions that can be drawn from the basic correlation matrix we can say

Table 1. Correlations between the variables of interest.

Variable	Experience	EPBS	ASGS-B	Trauma
EPBS	0.551	---	---	---
ASGS-B	0.316	0.373	---	---
Trauma	0.440	0.168	-0.051	---
Fantasy	0.272	0.271	0.180	0.131

that a fair degree of consistency has been evinced, a few errant correlations notwithstanding. This is encouraging given the nature of the sample, and perhaps warrants a little more confidence in subsequent interpretation of the data.

In the previous study a detailed tabulation of the specific traumatic correlates of belief and experience provided clear evidence of a stronger link between trauma and experience than was evident for trauma and belief; a result which was also suggested by the causal modelling results in that study. In table 2 I present similar results for the present study. Naturally this table differs in its inclusion of correlations for both belief measures. It also differs in two other respects. Firstly, in the published data for the previous study the means were composed of averages averaged by subject and by the number of items contributing to that subject average (e.g. there are three items for familial sexual abuse but only two for robbery/mugging). This item averaging is not present in the present table, accounting for the difference in means and standard deviations between the studies. If one includes the item averaging for the present data then there is a fair agreement between the means and standard deviation for these two studies, save for the isolation mean, which brings us to the second difference. After the previous study was published I noticed an error, committed by both Irwin and myself, in the calculation of the isolation sub-score. Whereas Irwin and I had both added the two items comprising the

Table 2. Source of Childhood Trauma	Mean	S.D.	Spearman rho (EPBS)	Spearman rho (ASGS-B)	Spearman rho (Experience)
Intrafamilial sexual abuse	3.034	0.263	0.210	-0.072	0.178
Intrafamilial physical abuse	11.190	5.886	0.171	-0.033	0.255
Loss (friend)	3.603	1.611	0.030	-0.033	0.046
Extrafamilial sexual abuse	3.552	1.062	-0.004	0.007	0.242
Loss (family)	4.414	1.487	0.066	-0.193	0.056
Isolation	3.138	3.663	0.222	0.027	0.439
Personal Illness	2.741	1.085	-0.180	-0.085	0.138
Parental divorce/miscarriage	2.241	0.506	0.177	0.011	0.344
Assault	3.034	1.426	0.179	-0.022	0.142
Loss of property (home moving)	4.759	1.302	-0.091	-0.177	-0.112
Robbery/mugging	1.138	0.437	0.166	0.054	0.097

isolation sub-score, the scoring system for the sub-scores indicates that the two items should be multiplied together. Thus if someone indicates three separate episodes of isolation of at least 1 months duration each on the isolation items this would score 3+3 on the incorrect scoring system and 3*3 on the new. The actual score would be 3 points lower than the correct score. This error would clearly influence both the mean and the standard deviation of the scores, and this is what is found in table 2. Because of the expansion in variability that would result from the correct scoring one would expect an increase in the strength of any real correlation that exists between isolation trauma and related variables due to the expansion of range.

Turning to table 2 then we can see that a similar pattern emerges as was found in the previous study. Spearman correlations between experience and the specific traumatic sources are significant for three variables, and near-significant for a further three variables. For the EPBS three variables show near-significant correlations, whilst only one variable shows any kind of near significant correlation with the ASGS-B. Once again then, we see that the correlations for the specific types of traumatic event are more salient and numerous when experience is the correlate than when either kind of belief

measure is the correlate, fully confirming the general findings of the previous study in this respect. Particularly noteworthy is the anticipated expansion of the correlation of isolation with experience. This is also felt somewhat for the EPBS, whilst the ASGS-B shows almost exactly the same degree of non-correlation as was found in the previous study. This result is very encouraging, as are the basic results from table 2 in general, providing once again subtle allusions to the model to come.

Results of Causal Modelling

Both models taken from the last study were separately analysed for their goodness of fit in this study on an a priori basis (i.e. no modification was made to the models prior to the analysis that might influence their degree of fit). Running EQS on the model given in Figure 1, the simplified version of Irwin's model, gave fit statistics as indicated in table 3, column 2. The fit statistics for Irwin's model indicate an inadequate degree of fit in all respects. The Satorra-Bentler chi-square is significant, indicating that the model leaves a significant amount of covariance left unexplained. The Bentler-Bonnett non-normed fit index is particularly poor, and this is problematic given its sensitivity to fit in small samples. The situation is somewhat improved for the confirmatory test of the Lawrence et al. (1995) model. Although on average that model leaves a slightly higher residue of covariation to be explained, this is due to just one or two marginally discrepant residuals[1]. Generally, the fit statistics for the Lawrence model are far improved, and this is reflected in the Satorra-Bentler scaled chi-square which is non-significant (indeed it is almost at its optimum value, equal to the degrees of freedom). Finally, the Bentler-Bonnett non-normed fit index is well within the acceptable 0.90 boundary for a well fitting

[1] Indeed, three residuals are above ±0.10; one of these being the, this time, non-significant relationship between trauma and fantasy. Two other relationships also contribute to the residuals; a negative one between truama and the ASGS-B, and a positive one between fantasy and the EPBS. Neither are marked out for inclusion in the LM test.

TABLE 3. MODEL FIT STATISTICS	Irwin's Model (fig.1)	Lawrence et al. Model (fig.2)
Average off-diagonal absolute standardised residuals	0.0639	0.0739
Independence Model Chi-Square (*df*)	51.67 (*10*)	
Independence AIC	31.67	31.67
Model AIC	6.46	-4.33
Satorra-Bentler Scaled Chi-Square (*df*)	15.65 (*4*)	6.86 (*5*)
Chi-Square *p*-value	0.0035	0.2309
Normal theory RLS Chi-Square	12.883	5.067
Bentler-Bonett Non-Normed Fit Index	0.372	0.968
ΔChi-square (*d.f.*) (*p-value*)	8.79 *(1) (p<0.005)*	

model. Despite the irritating failure to confirm a significant relationship between trauma and fantasy (generally not disputed in the psychological literature) it is clear that generally the Lawrence et al. (1995) model is confirmed in this small study. Thus, the relationship between trauma and fantasy notwithstanding, the results of the previous study were not merely the result of capitalisation on chance.

Whilst the Lawrence model reproduces the original correlations in the basic data very well, it might not necessarily do this with similar structural coefficients. Figure 2 shows the full path diagram for the model with the relevant components of the standardised measurement and structural models represented. From this model it is clear that the essential pattern of the structural coefficients is replicated, save for the effect of trauma upon fantasy. In particular the coefficient for the effect of fantasy on belief is exactly reproduced. The coefficient for the novel direct effect from trauma to experience is not only well confirmed, but is somewhat stronger in the latest modelling attempt. The coefficient for the effect of experience on belief is also

Table 4.

	PES	ASGS-B	EPBS
Expressed Control	0.241	0.035	0.197
Wanted Control	0.028	-0.158	-0.059
Expressed Inclusion	0.036	-0.088	0.151
Wanted Inclusion	0.084	-0.165	0.109
Expressed Affect	0.130	-0.077	0.041
Wanted Affect	0.002	-0.069	0.007

appropriately high. The appropriateness of this particular relationship is manifested in two respects. Firstly, one has come to anticipate high correlations between experience and belief, but secondly higher path strengths are a common feature of models using latent variables. Thus, with the diluting effects of systematic unique variance factored out of the experience-belief relationship one naturally obtains a more accurate estimate of the genuine relationship between the two. Paradoxically, some researchers regard this kind of effect enhancement with the suspicion that it is some kind of mathematical sleight of hand. Quite the opposite, assuming both factored variables do share a common core then the latent variable will be composed of this core, whilst unique but systematic variability specific to each measure is left out - this must surely be a valuable feature of the latent variable model! On the integrity of the measurement model of the belief latent variable we see that both belief measures adequately load the factor of paranormal belief.

Thus, despite the small size of the sample, and despite its apparent particularity, the results of this study fairly support those of the previous study.

Results for the FIRO-B scale

In table 4 we can see how the relationship between the six subscales and the three paranormally relevant scales breaks down. The relationship between the EPBS and expressed control (rho = 0.197) most clearly supports Irwin's earlier findings, where Tobacyk's PBS gave a similar moderate correlation with expressed control (rho = 0.32) when assessed by the survey respondents assistant, or a non-significant correlation (rho = 0.10) when the survey respondents themselves filled out the FIRO-B. Strangely, the ASGS-B shows no relationship with expressed control, and a non-significant but nevertheless weakly negative relationship with wanted control. Meanwhile, once again, the PES shows the strongest correlation with expressed control, and almost no relation to any other variable save for a small correlation with expressed affect. Clearly the data at hand tend to support the notion that expressed control alone has some part to play in future modelling attempts. I now turn to a discussion of these results.

Discussion

Rather than being the result of capitalisation on chance, the model suggested in the last chapter's specification search would appear to depict a fairly robust and consistent representation of the development of paranormal belief and experience, at least so far as the pattern of results for childhood fantasy and childhood trauma is concerned. The rather restricted nature of the sample, and its small size, has not prevented a similar pattern of results from presenting in this study. Despite this success, the SPR study proved too small to allow anything other than this most basic confirmation of the Lawrence et al. (1995) model. The larger model outlined earlier in this thesis could not be analysed due to empirical underidentification (the variable:sample size ratio was too low). Further still, the simplification inherent in the present and previous modelling attempts means that strictly speaking those previous models have been misspecified. With the inclusion of all, or most, of the relevant variables in Irwin's model, one may yet find significant changes to the parameters that

have been specified. Lastly, the analysis of the FIRO-B measure, though basic, has lent some more support to the notion that a need for control is manifest in the character of the paranormal believer and experiencer. The final study remedies these problems, and it is to this study and its results that I now turn.

Chapter 8

THE EDINBURGH SURVEY

We have now seen two successive attempts to confirm and extend Irwin's childhood factors model of paranormal belief and experience. Some interesting modifications were suggested in the first modelling attempt, and when the model based on these modifications was cross-validated on the SPR data the results provided a reasonably good degree of corroboration for the first study's respecification. However, more interesting issues came out of the first study that could not realistically be dealt with in the SPR study, primarily because the sample size for that study was too low. It remained then for me to attempt a full blown cross-validation of the previous study results on a larger, more representative sample. Likewise, my predictions about the meaning of the direct route from trauma to experience would need to be put to specific test. These are the key features I wish to develop in this chapter. If the model developed in previous chapters is seen to cross over to the wider population of people living in Edinburgh then we may feel fairly confident that a stable set of relationships is being addressed, something that is worthy of concerted theoretical analysis in the concluding parts of this thesis. Of course I cannot be sure that my model includes all the relevant factors, but in the absence of any compelling reason to believe the model omits key variables I would feel confident, if the previous model was confirmed, in offering up these results for serious consideration as a fundamental component in the understanding of paranormal belief and experience.

There are basically two planned analyses in this chapter. The first is the basic confirmation of the model developed in the previous two chapters, and as we are now familiar with that model I shall not dwell on it here. However, a larger study

should afford us the possibility of exploring in more detail some aspects of Irwin's more complex specification of the development of paranormal belief and experience. Secondly, in Lawrence et al. (1995) I proposed a two route model of the relationship between trauma and paranormal experience, in which trauma has both direct and indirect influences on paranormal experience. The indirect route is moderated by childhood fantasy, and represents the widely accepted view that childhood fantasy, in the form of fantasy proneness, has a prominent role in the development of paranormal experience. The direct route is a novel addition to the model proposed by Irwin, and seems warranted from the results of the second study. In brief, I argued earlier that the two routes may well stem from two kinds of trauma; control related and bereavement related. It seems the best way to test this hypothesis is to model the factors relevant to it on their own, divorced from complication with other factors. This then is what I shall aim to complete in this chapter. For clarity of exposition this chapter will be split in to two parts; one which provides the confirmation of the previous two chapters' findings, and the second, which tests what I shall call the 'specified trauma model' of traumatic influences on paranormal experience.

Methods

Subjects

The respondents were 129 citizens living within the city of Edinburgh, from an initial survey acceptance quota of 300. The volunteer response rate is therefore 43.66%, by no means a poor response rate for this kind of survey. However, 134 people simply refused outright to participate in the survey. A large majority (64%) of these were judged to be over 50, and as so many of these refused to participate *before* the survey could be explained to them, there is a somewhat lessened chance that strong content related selection biases apply. Of the 116 people who gave

their sex, 45 were male and 71 female. One hundred and fifteen people gave their age, and for this sample the mean and median age were 40 years.

Materials

A 225 item survey questionnaire was constructed using items from a variety of scales. Only the following are relevant to this study;

(1) Survey of Traumatic Childhood Experiences (STCE) -- This scale, developed by Council and Edwards (1987), comprises 30 items dealing with events that may cause trauma in a child's life. The full 30 item scale was used in this study. The scale has items relating to intrafamilial sexual and physical abuse, loss related to friends and family, extrafamilial sexual abuse, assault, loss of home and property, isolation and personal malady, adolescent abortion/miscarriage, and parental divorce/separation. Responses are scored on a 5 point scale of frequency, with (A) none, (B) one, (C) two to five, (D) six to ten, and (E) more than ten traumatic experiences.

(2) Modified Australian Sheep-Goat Scale (ASGS) -- this scale was a shortened version of Thalbourne and Delin's (1993) ASGS. Only those items referring to paranormal belief were used (the ASGS also refers to paranormal experience). Items were drastically abbreviated to make their meaning clearer on the advice of colleagues whose own participants voiced some confusion about the nature of the original wording for items in the ASGS.

(3) Edinburgh Paranormal Belief Scale (EPBS) -- this scale was drawn together by the author in order to overcome some of the other problems with other scales for the measurement of paranormal belief. The intention was not to produce a psychometrically impeccable instrument, but merely to have a second tool for the assessment of paranormal belief that might complement

the ASGS (thus cancelling out their individual deficits as measures of that construct).

(4) Childhood fantasy -- This construct was measured using only those items from Myers' (1983) ICMI:C that refer explicitly to childhood fantasy. Items are scaled dichotomously, in the true/false format. High scores indicate high fantasy.

(5) Paranormal Experiences Scale -- This scale measures the extent to which a person has experienced events of an ostensibly paranormal nature. Comprising 14 items, the scale measures the following kinds of paranormal experience; general psi ability, telepathy, precognition, visions (general and of spirits), psychokinesis, out-of-the-body experiences, and anomalous object movement.

Survey Procedure

The random function of a Casio calculator was used to select one page from the Edinburgh residential telephone book. The first thirty street names in the left most column of this page were selected at random from the telephone book for the Edinburgh district. The author then visited each street, and exhaustively sampled at least 10 houses until the survey questionnaire had been accepted by ten individuals, whereupon a new street was sampled (on occasions a street was exhausted before ten questionnaires had been administered, in which case the author moved in to an adjacent street). Overall the author visited 434 houses until all 300 questionnaires had been accepted for completion. People answering the door were (if over 18) invited to participate in a study to examine the causes and consequences of paranormal belief and experience. Attention was drawn at this stage to the sensitive nature of some of the scale items, and participants were informed that should they not wish to complete part or all of the questionnaire they were perfectly within their rights to do so. Any questions about the nature of the

questions were answered to the satisfaction of the respondent, and if a potential respondent declined to take part a brief apology for calling on them was offered and the researcher departed. Participants were given the survey questionnaire to fill in. Participants were asked to fill in the questionnaire as honestly as possible, and the questionnaire introduction also included an honesty statement to dissuade participants from giving only socially desirable responses (Eysenck, Eysenck and Shaw, 1974). Despite the intimate nature of some of the questionnaire items, no respondent reported having any particular problem with the content of the scales. Participants were given one week to complete the questionnaire before the author personally returned to collect the questionnaire. If people requested more time to complete the scale a further week was given.

Part I - Results

Preliminary analysis

Examination of univariate and multivariate kurtosis estimates revealed no substantial departure from univariate or multivariate normality (Mardia's[1] coefficient = 0.8801, z = 0.5974, p=NS). Examination of case numbers with the largest contribution to the normalised multivariate kurtosis revealed only one strongly discrepant case, case 78. Examination of raw data for case 78 revealed that this correspondent reported a higher than average amount of loss related to friends and family. However, this case did not contribute to overall significant kurtosis, and so was not removed from the analysis.

An intercorrelation matrix was analysed using maximum likelihood covariance structure analysis. The determinant of the input matrix was suitably non-zero, and the necessary program statement, 'PARAMETER ESTIMATES APPEAR IN ORDER, NO SPECIAL PROBLEMS WERE ENCOURNTERED DURING

[1] See Mardia (1970) or (1974).

OPTIMIZATION.' given. This being the case, we shall turn to a consideration of the modelling results.

Modelling results

Table 1. gives the modelling results for the covariance structure analysis of the Irwin and Lawrence et al. (1995) models of paranormal belief development. It is manifestly apparent that the twice suggested superiority of Lawrence's model is confirmed in this analysis on a larger, more general sample. The off diagonal residuals for both models are good. The model Aikaike Information Criterion chi-square indicates that Lawrence's model is again the more parsimonious of the two. Interestingly, chi-square goodness of fit tests for both models are non-significant indicating that both models reproduce the covariation amongst the manifest variables very well. The normal theory RLS chi-square supports this conclusion also. Finally, the Bentler-Bonett non-normed fit index gives a very acceptable degree of model fit for both models, favouring Lawrence et al. model. Despite the improved quality of fit for Irwin's model on this sample of data, it is now abundantly clear that the Lawrence et al. model represents a genuinely and consistently better fit to data collected using the variables of trauma, fantasy, paranormal belief and experience. Cautious use of the specification search facility of EQS by Lawrence et al. (1995) has resulted in a model which cross-validates reasonably well in widely differing samples. However, only examination of the largest standardised residuals, standardised solution results, and Lagrange and Wald tests, will indicate whether the model's quality of fit is precisely the same as for previous analyses.

For Irwin's model the three largest residuals were for the correlations between trauma and the EPBS (0.176), trauma and paranormal experience (0.119), and trauma and the ASGS-B (0.107). All other residuals were within ±0.004 points of zero error. Once again, for Irwin's model the relationship between trauma and

TABLE 1. MODEL FIT STATISTICS	Irwin's Model (Fig.1)	New Model (from Lawrence et al., 1995) (Fig.2)
Average off-diagonal absolute standardised residuals	0.0406	0.0422
Independence Model Chi-Square *(df)*	221.064 *(10)*	
Independence AIC	201.0639	
Model AIC	-2.97	-4.93
Chi-Square *(df)*	5.028 *(4)*	5.069 *(5)*
Chi-Square p-value	0.28443	0.40750
Normal theory RLS Chi-Square	4.928	4.927
Bentler-Bonett Non-Normed Fit Index	0.988	0.999

paranormal experience is found to be poorly explained. All measurement and construct equations had highly significant coefficients in the model, except for the path coefficient representing the influence of fantasy upon paranormal belief (coefficient = 0.041, z=0.203, p=NS). This of course hints at the likely removal of that link from the model in the forthcoming Wald test, a result in keeping with the

two previous studies results. True to form, the very first parameter considered for removal in the Wald test is the link from fantasy to belief. For the Lagrange multiplier results, which suggest possible additions to the model, only one suggestion gives a significant chi-square reduction; the addition of a link from belief *back* to fantasy. I find it difficult to make this link theoretically meaningful, especially as the path coefficient for this link would be negative (-0.264)! I shall therefore blatantly ignore this modification, passing it off as a candidate for capitalisation on chance if seriously implemented in the model. The very next suggested addition, which is not quite significant, is the familiar addition of a direct link from trauma to paranormal experience (chi-square = 2.55, p=0.110). Although not significant, the estimated parameter change resulting from adding this link is the largest of all those suggested (the fact that it is insignificant would suggest that

the standard error of the estimate is quite large). Nevertheless, it is clear that the direct link from trauma to experience is here to stay.

Turning to the largest standardised residuals for Lawrence et al. model, the largest three are, in order, the residual for the relationship between trauma and the EPBS (0.179), trauma and paranormal experience (0.119), and trauma and the ASGS-B (0.109). The remaining residuals are all within ±0.004 of zero error. The residuals are all too small to worry over, and indeed the Lagrange multiplier test for adding parameters found no unspecified parameters which if added would significantly enhance model fit. The Wald test did suggest that the direct link between trauma and experience be dropped (univariate $\chi^2_{(1)} = 3.097$, $p<0.078$, and so this path was removed and the Lawrence model reanalysed. The resulting model, with standardised path coefficients, is depicted in Figure 1.

Part 1 Discussion

The results of this confirmatory analysis are reasonably clear. The new model of paranormal belief development resulting from Lawrence et al. (1995) specification search is generally a better fit than Irwin's original model. There has been some difficulty with replicating the links between trauma and fantasy (in the last chapter), and trauma and paranormal experience (this chapter). However, even paths which have failed to be significant in confirmatory tests have only just failed to be significant. To be fair, we are still dealing with smallish sample sizes in this chapter, with all that implies for replication of weaker links. However, Lawrence et al. suggested that the two routes to experience from trauma might have different traumatic causes. I shall not herein go over the details, suffice to say that the next part of this chapter aims to resolve the status of this conjecture, which at present has only conceptual appeal.

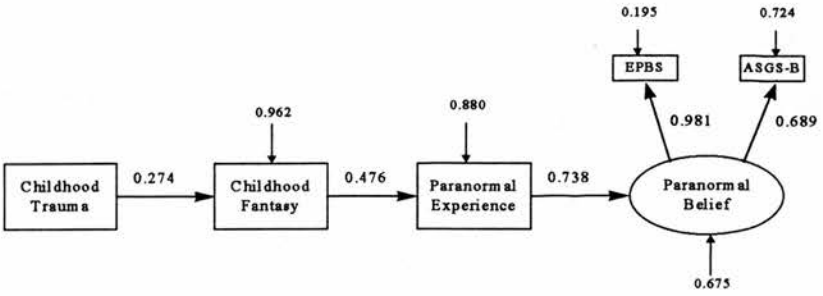

Figure 1. The full path diagram for Lawrence's model of the childhood factors minus non-significant paths.

Part 2 - Examining the Specified Trauma Model

In chapter 6 I attempted to explain the evidence for a direct link from trauma to experience by proposing that the direct route is caused by a specific type of trauma, bereavement related trauma. In contrast, the indirect effect of trauma on paranormal experience is supposed to be the product of control related trauma, the kind of trauma explicitly theorised in Irwin's (1992) paper on the origins of paranormal belief. In this section of the chapter I shall directly test this hypothesis.

The first consideration is just how to operationalise the two traumatic latent variables of control and loss related trauma. In the one instance we want a loss factor indicated, at a minimum, by variables relating to loss of friends and family

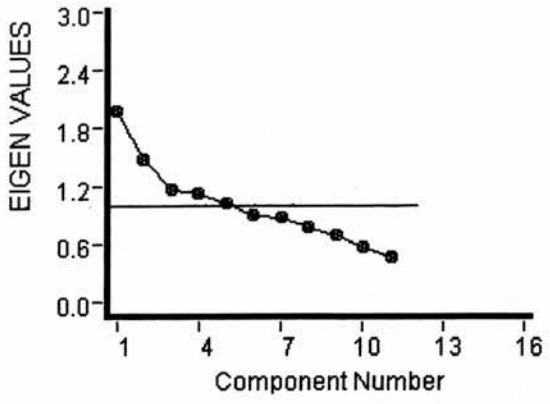

Figure 2. A scree plot of eigenvalues for the factor analysis of the STCE.

on the STCE, whilst the control factor should relate together traumatic variables like family physical abuse, parental divorce, and loss of home at the least. One possible solution would be to perform an exploratory factor analysis on the STCE main variables to see a) what the dimensionality of the STCE is, and b) whether within the factors selected there is clear evidence of control and loss factors. It just so happens that the results of a principal factor analysis of the 11 main variables of the STCE do indeed support the notion that there are two factors of trauma, one relating to loss traumas, and one relating to control traumas. The results of this analysis are given in table 2 and Figure 2.

Initially a principal components analysis was performed on the STCE, giving the scree plot in Figure 2. From this we see clear evidence for two genuine factors of the STCE. This is indicated by the marked increase in variance accounted for after factors 3-11. From factors 3-11 there is a very constant and tiny increase in variability accounted for by each factor - as one would find with a factor analysis

Table 2. Factor Loadings for the Principal Factor Analysis of the STCE (Orthosim Solution)

Variable	Factor 1	Factor 2
Intra-familial Sexual Abuse	-0.0742	0.0267
Intra-familial Physical Abuse	0.0572	**0.3845**
Loss of friends	**0.5255**	0.1406
Extra-familial sexual abuse	0.2577	0.2651
Loss of family	**0.5196**	0.0780
Isolation	**0.5052**	0.0104
Illness	**0.3420**	-0.0779
Parental Divorce	0.0548	**0.4348**
Assault	-0.0378	0.2791
Loss of home	0.0931	**0.3635**
Robbery	-0.0557	**0.3287**

on random variables. But the 'mountain' of genuine factors rises above this general trend only with factor 2, the rest being mere scree. So two factors were selected for analysis and rotation using principal factor analysis. Both orthosim (orthogonal) and direct oblimin (oblique) rotations were performed, but as the correlation between factors in the oblimin solution was a mere 0.074 the orthogonal solution was interpreted. The factor loadings for the orthosim solution are given in table 2. Taking as a minimum loadings greater than 0.30, we see that factor one is loaded on by loss related to friends, loss related to family, isolation, and illness. Here then is our loss factor. Likewise, factor 2 is loaded on by family physical abuse, parental divorce, loss of home, and robbery. The two most strongly loading variables for this factor are clearly of control-theoretic relevance, and a similar though less strong case can be made for the loss of home and robbery variables. In addition simple structure appears to have been achieved, excepting the loading for extra-familial sexual abuse. The almost zero loadings for intra-familial sexual abuse merely reflect the fact that there practically is no variability to be explained here.

From this factor analysis we have come away with no little support for the two kinds of trauma envisaged by Lawrence et al. (1995). Eight of the eleven variables

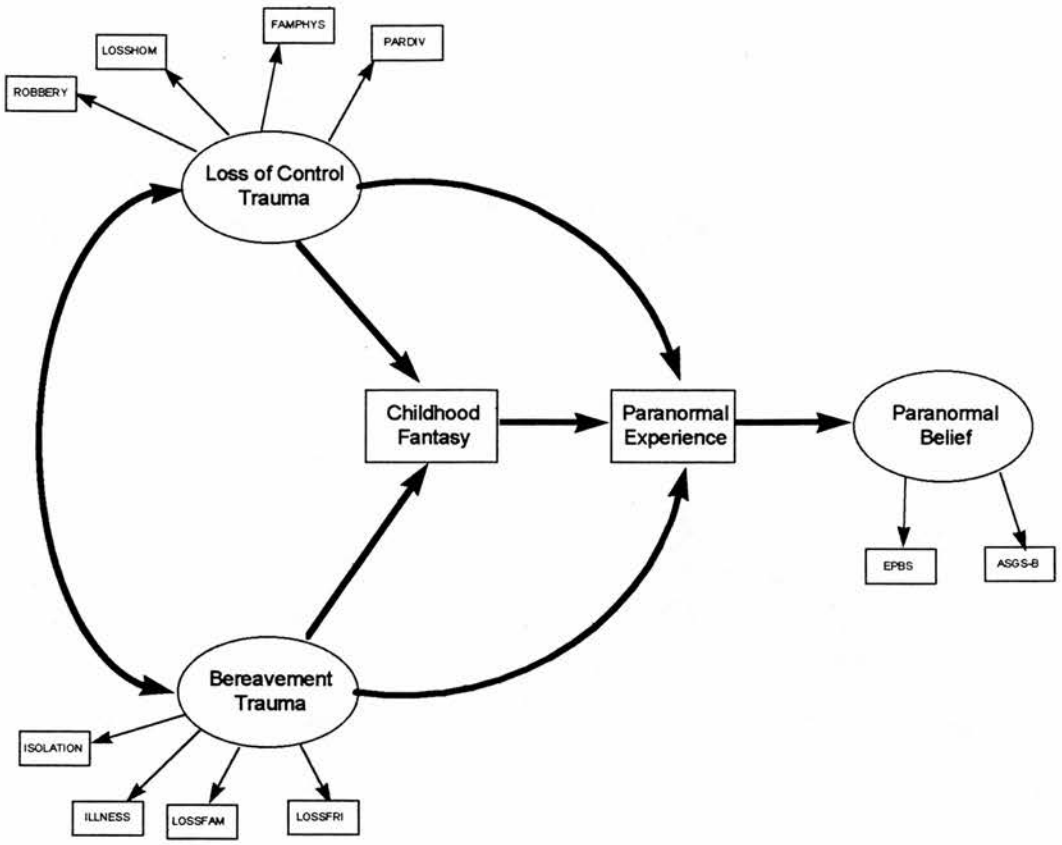

Figure 3. The Specified Trauma Model

will go on to play a part in the test of the specified trauma model, and it would seem appropriate to use the factor loadings from this exploratory analysis as the start values for the latent constructs of loss and control related trauma in the subsequent analysis.

We now must consider the initial path diagram to test the specific trauma model of Lawrence et al. (1995). Figure 3 provides a path diagram that precisely depicts Lawrence et al. hypotheses about the two routes from trauma to paranormal experience. As we can see the two factors of trauma, control and bereavement

related, are indicated by the variables based on the principal factor analysis of the STCE reported earlier. They are both permitted to correlate in the initial test. With both factors initially 'causing' both fantasy and experience we have our best chance of determining if Lawrence et al. are right. If with this model we find the route from control related trauma to fantasy to be stronger than that from loss related trauma to fantasy then we have evidence that control related trauma is the main cause of fantasy. Likewise if the direct route from loss related trauma to experience is greater than that from control related trauma then we have support for the notion that loss related trauma is responsible for the direct route to paranormal experience. Furthermore, my conjecture would have its strongest confirmation if the path coefficients for the direct route from control trauma to experience and the direct route from bereavement trauma to fantasy are estimated at or near zero. Such a result would indicate the absolute separation of these factors effects.

Lastly, for this modelling attempt I should note that, apart from those items used to fix the scale for each factor, the start values for the items loading on to control and bereavement related trauma are all set at the level of their exploratory factor analytic loadings.

Part 2 - Results

Preliminary analyses

Examination of univariate and multivariate kurtosis estimates revealed substantial departures from both univariate normality and multivariate normality (Mardia's coefficient = 36.99, z = 11.46, $p<0.01$). Univariate estimates of kurtosis and skew were outside acceptable ranges for loss of friends, loss of family, isolation, parental divorce, and robbery. Examination of the largest contributors to multivariate kurtosis revealed that, once again, case 78 was causing the severe departure from non-normality, aided by a less discrepant contribution from case 70. In order to combat any problems in statistical inference that may result from this non-

TABLE 3. MODEL FIT STATISTICS	Specified Trauma Model (Fig. 4)
Average off-diagonal absolute standardised residuals	0.0518
Independence Model Chi-Square *(df)*	309.838 *(45)*
Independence AIC	219.838
Model AIC	-34.046
Chi-Square *(df)*	32.953 *(32)*
Chi-Square p-value	0.42026
Normal theory RLS Chi-Square	32.697
Satorra-Bentler Scaled Chi-Square	**33.4489**
Bentler-Bonett Non-Normed Fit Index	0.995

normality, the subsequent analysis will be run with robust statistical estimators (see Satorra and Bentler, 1988a, 1988b, and Hu, Bentler and Kano, 1992).

Thus, an intercorrelation matrix was analysed using maximum likelihood covariance structure analysis with robust statistical estimators. The determinant of the input matrix was non-zero[2], and the necessary program statement indicating no special problems in optimisation was issued.

Modelling Results

The model as specified in figure 3 provided a very good fit to the data, but contained a number of clearly non-significant parameters. In particular, the paths representing the effect of loss of control trauma on paranormal experience, and of beareavement related trauma on childhood fantasy were non-significant as I had predicted. In addition, the loading of robbery on loss of control trauma was non-significant, as was the loading of illness on bereavement related trauma. Finally, the effect of bereavement related trauma on paranormal experience was, although strictly non-significant, so very close to significance that it was retained in the respecified analysis below. With all but one non-significant parameter removed

[2] A non-zero determinant is needed to calculate the inverse of a matrix, which is needed to perform the matrix equivalent of division. When the determinant is zero the matrix is said to be singular and division by zero occurs, bringing all computation to a halt.

from figure 3, the respecified model was analysed (minus non-significant parameters). Table 3 gives the modelling results for the respecified trauma model as specified in Figure 4. We can see that with average off-diagonal residuals of 0.0518 the model as finally specified turns in a good degree of fit between sample and estimated correlation matrices. The likelihood ratio chi-square of 32.953 indicates a non-significant degree of discrepancy between model estimate and sample data; indeed, chi is almost at its asymptotic optimal value, equivalent to the degrees of freedom. The normal theory RLS chi-square also supports the standard chi-square test result. Most importantly, the scaled chi-square test statistic, which is robust to violations of multivariate normality, is well within the bounds of non-significance. This suggests that the results can meaningfully be interpreted as representing a well fitting model. Lastly, the Bentler-Bonett non-normed fit index is very close to optimal at 0.995, and well within the bounds of acceptability. In short all results indicate that the respecified trauma model provides an excellent fit to the sample data.

Examination of the largest standardised residuals shows that the respecified model fails to account adequately for the relationship between variables isolation and fantasy (residual = 0.185), and family physical abuse and the EPBS (residual = 0.180), these being the two largest residuals. All other residuals are of marginal size, within ±0.01 of zero error.

With these caveats in mind the standardised solution path diagram, with error coefficients for dependent variables, is given in Figure 4. Despite the poor quality of the traumatic factors the structural component of the model bears a close resemblance to Lawrence et al. (1995) hypothesis. Control related trauma, insofar as it is identified with family physical abuse and parental divorce, is emphatically

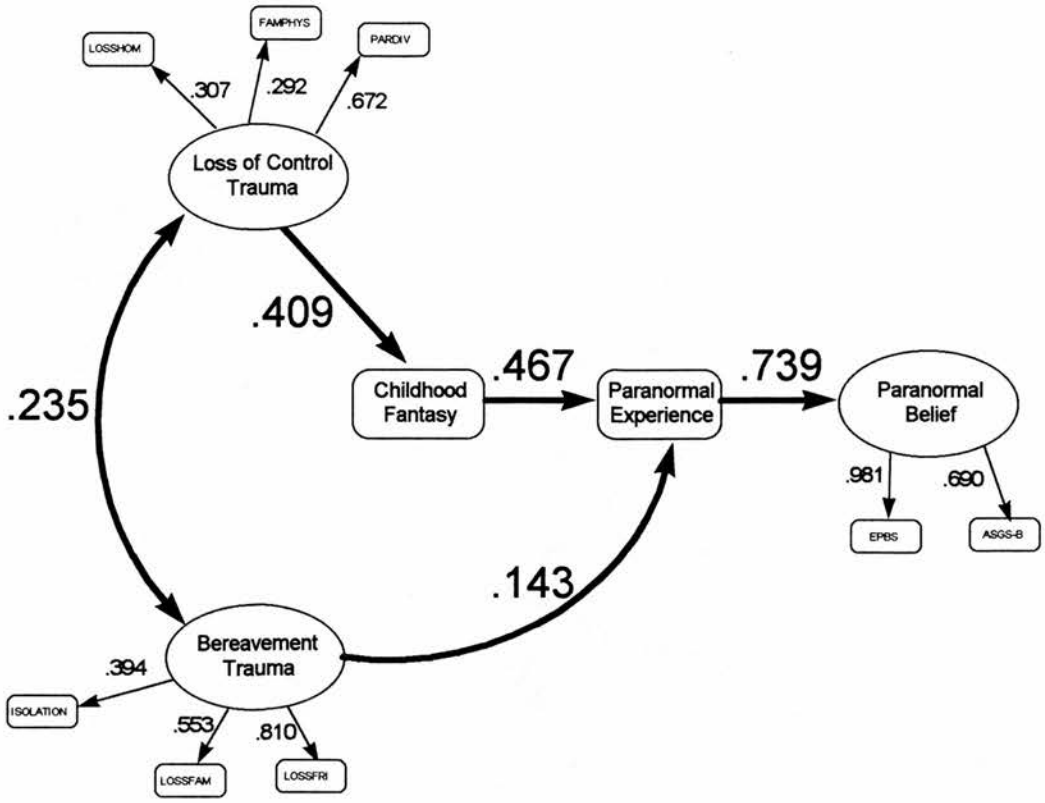

Figure 4. Final Specified trauma model with results, minus non-signficant parameters.

related to fantasy proneness, but not to paranormal experience. Likewise, the only evidence, albeit of marginal significance, for an effect of bereavement related trauma within the model, shows that its primary influence is a weak direct effect upon paranormal experience, with no effect upon fantasy proneness. Thus, though the overall result is not emphatic, Lawrence et al. (1995) conjecture about the direct route from trauma to experience seems to be supported.

Discussion

The analyses conducted in this chapter broadly lend further weight to my own model of paranormal belief development. This model continues to better Irwin's

(1992) original model of the childhood causes of paranormal belief and experience, suggesting that it genuinely captures some true feature of paranormal experience development. Further still, and perhaps more impressive, I proposed a specific cause for the two routes to paranormal experience in their new model, and the attempt to confirm my hypotheses about these two routes met with considerable success, despite some inadequacies in the measurement of the constructs of control and bereavement related trauma. Nevertheless, this chapter has taken the original models proposed by Irwin (1992) and myself a step further. The final chapter on the theme of modelling attempts to address the wider model proposed by Irwin. In particular I shall attempt to address the question of whether paranormal belief once developed actually has consequences for the individual, in terms of meaning in life and well being.

Chapter 9

PARANORMAL BELIEF, MEANING IN LIFE AND WELL BEING

In this chapter I wish to examine the role that paranormal belief and experience may play in the maintenance of meaning in life and well being. This can be considered as an examination of the tail end of Irwin's (1993) general model of paranormal belief development. According to Irwin (1993), paranormal beliefs and experiences foster an illusion of control, and this in turn leads to enhanced coping with uncontrollable life events. If this is indeed the case then we might suppose that people with high paranormal beliefs should show enhanced levels of meaning in life and psychological well being, in contrast to those with low levels of paranormal belief. However, this is not necessarily the case. It could be that those who experience a lot of trauma have very low meaning in life and well being (initially), but that paranormal beliefs serve to bring them up to only average levels of coping. Analyses of correlations, as form the basis of this chapter, will not enable us to decide between these two possibilities, but merely whether an illusion of control (somehow defined) covaries with meaning and well being.

The difficulty in testing Irwin's general model is in working out how to operationalise such constructs as 'illusion of control', or 'coping with uncontrollable life events'. Illusion of control, as a term, has a definite meaning in experimental work relating paranormal belief to success in parapsychology experiments testing for psychokinetic ability (Blackmore and Troscianko, 1985). In those experiments an illusion of control is said to have been demonstrated when the participants' judgement of success in the task at hand exceeds their actual level of success. However, in the real world it is nigh on impossible to determine whether one suffers from an *illusion* of control,

excepting cases of known psychopathology. To do so would require that one knew just how much control a person actually exerted over events in their life, and how much control was exerted over them by those events. As the real world is massively multivariate it would be unrealistic to attempt to measure any illusion of control.

It may however be entirely possible to examine the extent to which people have an internal or external locus of control. Here we are asking merely whether the person feels controlled by, or controls, external variables (i.e. a phenomenological issue not an ontological one). On this front there is a reasonable body of work which suggests that paranormal belief is associated with varying levels of internal and external control. However, the picture here is quite complicated, and deserves extended examination.

Locus of control and paranormal belief

Locus of control (Rotter, 1966), as a construct, refers to the extent to which a person feels that they are in control of events or events are in control of them, where 'events' could be manifest at the level of world politics, or office politics, or domestic politics, or even personal behaviour. In contemporary psychology there is a proliferation of measures of locus of control (Furnham and Steele, 1993), leading some to suppose that locus of control is a barren concept with no explanatory power (Kline, 1992, is an example). Nevertheless, at a basic level people do seem to differ in the extent to which they express an internal or external locus of control, and these differences often covary with a number of psychological variables. Interest in the relationship between locus of control and paranormal belief stems from the work of Adorno, Frenkel-Brunswick, Levinson and Sanford (1950) who suggested that superstitious beliefs were strongly related to belief in supernatural external influences upon the person. In keeping with this, Jahoda (1970), Scheidt (1973), and Jones et al. (1977) all found a significant correlation between belief in the supernatural and external locus of control. However, these early studies depended upon a somewhat

restricted conception of paranormal belief (that it was all supernatural or all just superstition). Fortunately, though it may sound odd coming from me after the reader has ploughed through chapters four, five and six, more recent studies of locus of control and paranormal belief have been able to exploit the more realistic multidimensional conception of belief introduced by Tobacyk and Milford's (1983) Paranormal Belief Scale. I shall exemplify this with reference to the most enlightening of all studies on locus of control, that of Davies and Kirby (1985).

Davies and Kirby sought to examine the relationship between perceived control and belief in the paranormal insofar as it is measured multidimensionally by the PBS. To do this they administered Paulhus' (1983) Spheres of Control Scale (SOC) and Tobacyk and Milford's (1983) Paranormal Belief Scale to 54 female and 41 male liberal arts students. The SOC is an interesting scale which measures three distinct types of control; socio-political[1], interpersonal[2], and personal control[3]. Likewise, my own critique aside, nominally the PBS measures seven factors of paranormal belief. Davies and Kirby were interested to see if the positive relationship between external locus of control would hold up across different types of paranormal belief, and different types of control (Rotter's I-E Scale correlates most strongly with the socio-political control dimension of the SOC - not studied here). In a canonical correlation analysis they discover two canonical variates. The first loads negatively upon the personal and socio-political dimensions of the SOC, and not at all upon the interpersonal dimension, whilst the second loads strongly and positively on the personal and interpersonal dimensions, and moderately negatively on the socio-political dimension. More interesting is that these two canonical variates have quite different loadings on the seven scales of paranormal belief. Thus, variate 1 had high loadings on religious belief,

---

1 E.g. The average citizen can have an influence on government decisions.

2 E.g. I have no trouble in making and keeping friends.

3 E.g. I can usually achieve what I want when I work hard for it.

superstition, and spiritualism, whilst variate 2 had high loadings on psi belief, witchcraft, and moderately on extra-ordinary lifeforms.

Thus, Davies and Kirby seem to have good evidence for an *internal* locus of control, at least in the personal and interpersonal realms, for those who express a high belief in psi phenomena of the kinds which concern us in this thesis. It is only more marginal belief factors, like religious belief, *superstition*, and spiritualism which are strongly related to external locus of control. This is of no little relevance for this thesis. If paranormal belief can assist in the establishment of an internal locus of control in the personal and interpersonal domains then it can have buffering effects upon any challenge to real control in those domains. And it is precisely those domains which are challenged when children or adolescents experience traumatic events of the kind considered previously in this thesis. Thus we might well expect interpersonal and personal locus of control to have effects upon purpose in life and directly or indirectly upon well being.

We should, however, be aware of the results of an attempt to replicate the Davies and Kirby (1985) study, conducted by Tobacyk, Nagot and Miller (1988)[4]. Though not using canonical correlation analysis, they did not find the same pattern of results at all in their southern US sample. Perhaps there are cultural differences going on here, with the largely Southern US population in Tobacyk at al. (1988) study having a genuinely different style of responding to control issues.

[4] Oddly, Tobacyk et al. (1988) make no reference to Davies and Kirby (1985) although the later study almost exactly addresses the same issue.

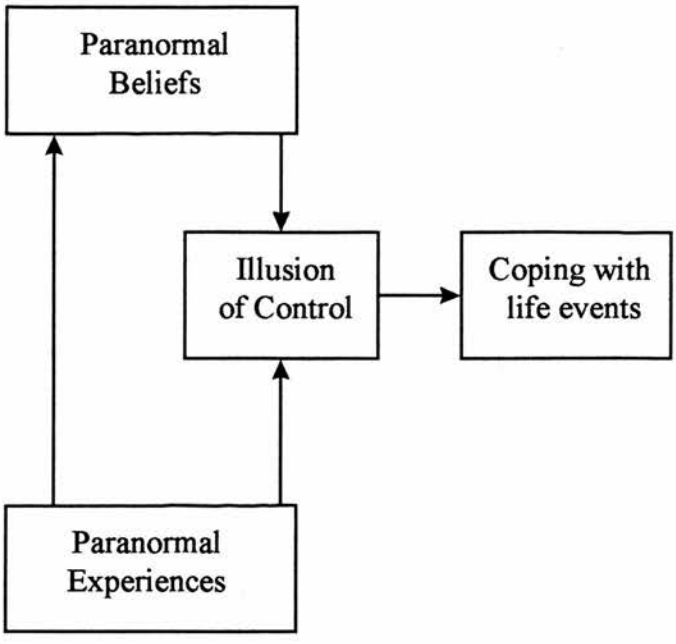

Figure 1. Irwin's (1993) model of the consequences of paranormal belief and experience.

This discussion now needs to be placed in the context of a preliminary model of the personal consequences of paranormal belief and experience. The immediate inspiration for such a model rests with Irwin's (1993) general conception of the consequences of paranormal belief and experience. In Figure 1 I depict this model as it appears in Irwin's original review of paranormal belief. I have, however, taken some licence with the model. For example, there is a causal link from paranormal experience to paranormal belief in the model as depicted here; Irwin's (1993) original specification was a mere correlation between the two. Also, Irwin has coping with life events[5] feeding back on to paranormal beliefs, reinforcing the buffering effect that they might have, but this is not the case for paranormal experience. I do not model this feedback effect, though it can be done. Such effects work less well in cross-sectional

[5] Irwin (1993) does not distinguish between genuine coping and illusory coping (i.e. the mere impression that one is coping). Presumably both kinds of coping may result from an illusion of control.

designs such as mine, and are best suited to longitudinal designs or three stage panel designs where the effects of x1 on y2 and y2 on x3 can be measured through time. Instead, we shall leave it to the Lagrange multiplier test to suggest such subtleties if needed. Using Irwin's model in Figure 1 one can, using the variables to hand in the Edinburgh survey questionnaire, form the following Irwin inspired model, which I shall call the personal consequences model. This is given in Figure 2. Though not immediately familiar, this model captures the essence of Irwin's specification as best as possible given the variables at hand. Thus, as both personal control and interpersonal control are to represent the illusion of control construct, and as the illusion of control construct is supposed to be caused by paranormal belief and paranormal experience, I have four causal links representing the effects of paranormal belief and experience on these variables. As I represent the construct of *Coping* with reference to two causally related variables, purpose in life and well being, the model differs here also. I hypothesise that the effect of a heightened sense of internal personal and interpersonal control is to enhance directly purpose in life and only through this improve well being. One could also start with a model in which the control variables are directly related to well being, but I think it best to start with a simple model and only build complexity if necessary. This then is the initial model designed to represent the personal consequences of paranormal belief and experience. I now turn to its testing.

Methods

Subjects

This study used the survey data taken from the Edinburgh survey, and the reader is advised to consult chapter eight for details of subjects.

Materials

From the Edinburgh survey questionnaire data, only the following variables are relevant to this study;

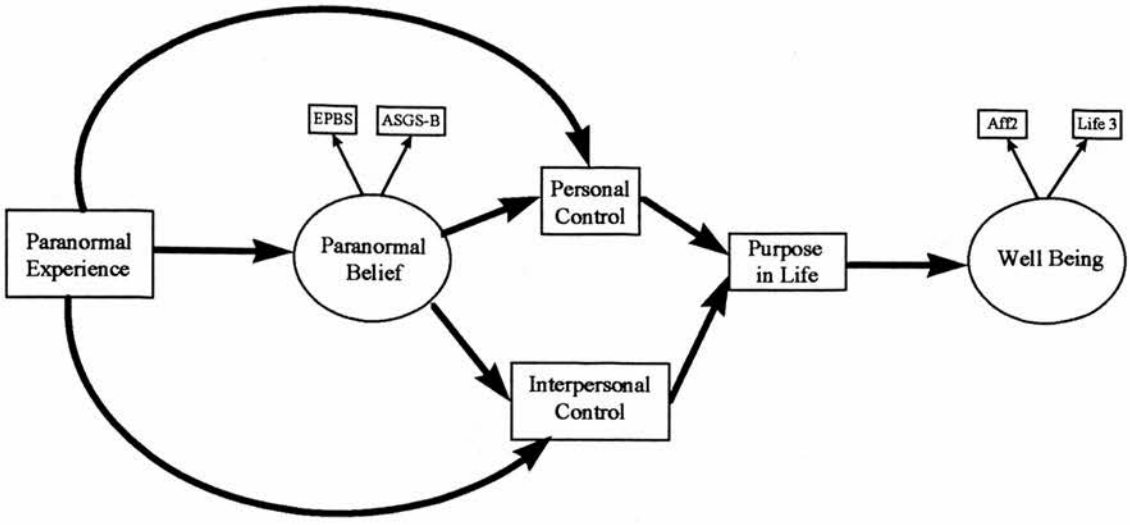

Figure 2. The Irwin inspired Personal Consequences Model Mk I.

(1) Meaning in Life – This was measured using the 20 item Purpose in Life (PIL) scale (Crumbaugh, 1968; Crumbaugh and Maholick, 1964). This test is designed to measure the degree to which a person experiences a sense of meaning and purpose in their life. Split-half reliabilities for the test have been reported in the region of 0.90 (Crumbaugh, 1968; Crumbaugh and Maholick, 1964; Reker, 1977), and test-retest coefficients in excess of 0.68 (Meier and Edwards, 1974; Reker, 1977). Zika and Chamberlain (1992) report an alpha coefficient for the scale of 0.91. Chamberlain and Zika (1988) recommend the use of the PIL as '...a general measure of meaning in life...' (p. 595).

(2) Modified Australian Sheep-Goat Scale (ASGS) -- this scale was a shortened 6 item version of Thalbourne and Delin's (1993) ASGS. Only those items referring to paranormal belief were used (the ASGS also unfortunately refers to paranormal experience). Items were drastically abbreviated to make their meaning clearer on the advice of colleagues

whose own participants voiced some confusion about the nature of the original wording for items in the ASGS. Thalbourne and Delin (1993) report reasonable factorial validity for a three factor structure of their original 18 item visual analogue scale. Internal consistency by Cronbach's alpha is 0.94, item-total correlations for the 18 items range from 0.45 to 0.79, whilst 11 month test-retest reliability is 0.66. See appendix for further details.

(3) Edinburgh Paranormal Belief Scale (EPBS) -- this 9 item scale was drawn together by the author in order to overcome some of the other problems with other scales for the measurement of paranormal belief. The intention was not to produce a psychometrically impeccable instrument, but merely to have a second tool for the assessment of paranormal belief that might complement the ASGS (thus cancelling out their individual deficits as measures of that construct).

(4) Well Being – this construct was measured using two published scales, Life 3 (Andrews and Withey, 1976) and Affectometer 2 (Kammann and Flett, 1983). Life 3 is reported to have a validity coefficient of 0.8 and a reliability of 0.7 (Andrews and Withey. 1976). Affectometer 2 is a 40 item scale with items measured on a 5 point proportion-of-time scale. It has an alpha of 0.95 and a 15 week stability coefficient of 0.72. Affectometer 2 measures both positive and negative affects (20 items on each), and in a recent study was found to correlate 0.63 with Life 3 (for positive affect) and -0.48 with Life 3 (for negative affect) (Zika and Chamberlain, 1992).

(5) Paranormal Experiences Scale -- This scale measures the extent to which a person has experienced events of an ostensibly paranormal nature. Comprising 14 items, the scale measures the following kinds of paranormal experience; general psi ability (including PK ability, precognition, and telepathy), abnormal visions, OBE's, apparitional experiences, and anomalous loss/discovery of objects. A factor analysis

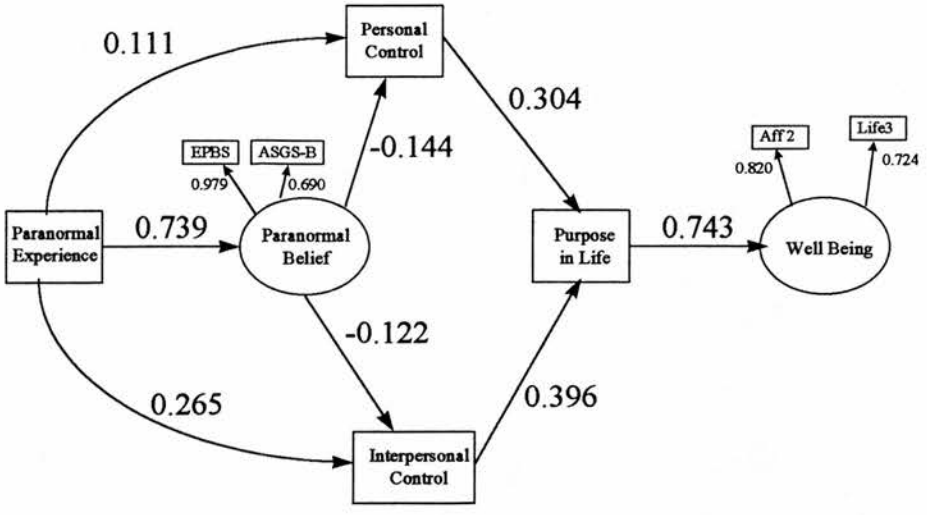

Figure 3. The Personal Consequences Model MK I with path coefficients.

shows that the scale is unidimensional, and all items have acceptable item-total correlations. It is scored on a 7 point scale of agreement.

Survey Procedure

The reader is referred to chapter eight for details of the survey procedure for this study.

Results

Preliminary Analyses

Examination of univariate and multivariate kurtosis estimates revealed no large deviations from normality, multivariate or otherwise. Mardia's coefficient of multivariate kurtosis = -1.46 (z=-0.66, p=NS). Examination of cases with the

largest contribution to multivariate kurtosis revealed no substantial outliers. This being the case the data were analysed using maximum likelihood covariance structure analysis.

Modelling results

Initial attempts at model analysis were thwarted by convergence problems[6]. In the first analysis with the model as specified in Figure 2 matrix estimation difficulties occurred from iteration 2, 5 and (twice) 7. In addition the parameter estimates for personal control and belief, and interpersonal control and belief were linearly dependent. Various attempts at solving the problem were tried, including fixing certain start values (factor loadings for paranormal belief were fixed at levels given in chapter eight results, along with the path coefficient between belief and experience) and increasing the number of iterations (to 100), but to no avail. However, Bentler (1993) suggested that in models where over-identification had been achieved, and where the program informs the user that parameters are linearly dependent, then, '...increasing TOL to, say, 0.001, may create a cleanly converged solution and sometimes also a dramatic reduction in the chi-square statistic.' (p. 71). As the model was over-identified with 16 degrees of freedom, the original model was analysed with the TOLERANCE[7] variable set to 0.001 instead of its default value of 0.000001. This produced a neatly converged solution in only 11 iterations, with no problems in parameter estimation, or optimisation difficulties.

---

[6] All successful modelling attempts with CSM end when 'convergence' has been achieved. For most typical applications of CSM convergence occurs when the fit function which EQS tries to minimise differs only at the 5th decimal place from the last iteration's fit function. EQS will continue to resolve the model until convergence is achieved, or the pre-set number of iterations is completed (usually 30). If EQS fails to converge before the pre-set limit is reached it issues a warning in its results output.

[7] According to Bentler (1993), 'Tolerance is a technical concept associated with pivoting operations in solving linear equations. It determines the point at which one variable can be considered to be linearly dependent on a set of other variables.' (p. 71). Frankly, I'm relying on Bentler's (1993) better judgement here.

TABLE 1. MODEL FIT STATISTICS	Personal Consequences Model Mk I	Personal Consequences Model MkII
Average off-diagonal absolute standardised residuals	0.0558	0.0615
Independence Model Chi-Square *(df)*	393.153 *(28)*	
Independence AIC	337.153	
Model AIC	2.904	-1.232
Chi-Square *(df)*	34.904 *(16)*	36.768 *(18)*
Chi-Square p-value	0.00410	0.00848
Normal theory RLS Chi-Square	32.694	34.159
Bentler-Bonett Non-Normed Fit Index	0.909	0.928

The results for this test of the 'Personal Consequences Model Mk I' are given in table 1 and Figure 3 (with path coefficients)[8]. For such a complex model, the fit statistics are surprisingly close to acceptable. Thus, the residuals are fairly small at only 0.0558. Chi-square indicates that the estimated correlation matrix based on this model fails to fit the observed correlation matrix. This is supported by the RLS chi-square also. Finally, the NNFI indicates only a close to marginal degree of fit against conventional levels of incremental fit. In short we might not expect to do too much respecification to find a model that fits (theoretical caveats excepting).

In assessing model misfit, the largest standardised residuals indicate that the correlation between variables personal control and interpersonal control, and interpersonal control and well being were poorly accounted for in the model. All other residuals were normally distributed around zero. Turning to the Wald test results only three parameters in the model are multivariately non-significant. These represent, in order of decreasing univariate p-value, the error term for the loading of the EPBS on belief, the influence of paranormal

[8] For clarity, error terms have been omitted from this path diagram. However, readers interested in recreating them can do so using the following formula;

$$error = \sqrt{1-\sum r_i^2}$$

Where r = the *i*th path coefficient causally effecting the dependent variable of interest.

experience upon personal control, and finally the influence of paranormal belief upon interpersonal control. What is of immediate interest here is that a) paranormal belief is slightly negatively related to both personal and interpersonal control, and b) the effect of paranormal belief on interpersonal control is small enough to warrant discarding them from the model. Secondly, paranormal experience is moderately related to interpersonal control, but only slightly positively to personal control. Indeed the relationship between experience and personal control is itself not significant on the univariate and multivariate Wald tests. Thus the only links between the paranormal variables and control variables are between paranormal experience and interpersonal control and paranormal belief and personal control, but the former link does at least fit with the work of Davies and Kirby in demonstrating that such perceived control tends to be internal control.

Of psychological interest is the discovery that both personal and interpersonal control are positively and moderately related to purpose in life (indicating that those who tend towards more internal control show higher purpose in life). Finally, the characteristically large correlation between purpose in life and well being (very well measured by its indicator variables) is apparent[9].

Turning to possible model additions, the Lagrange multiplier test suggested only two significant additions to the model, though both of these would greatly enhance the fit of the model. The first one was to allow for a link from well being back to interpersonal control, whilst the second was to allow for a link from well being back to personal control. These links are potentially theoretically meaningful (at least for psychology) as they imply that well being reinforces a sense of control over the personal and interpersonal domains. However, the data to hand are not longitudinal and thus cannot support the kind of effect suggested by the Lagrange Multiplier test, which would in this model lead to a backward loop. Thus, though the suggested modifications from

[9] Zika and Chamberlain (1992) report that the correlation between the PIL, and Life 3 and the positive dimension of Affectometer 2, is 0.71 and 0.78 respectively (based on a sample of 176 mothers).

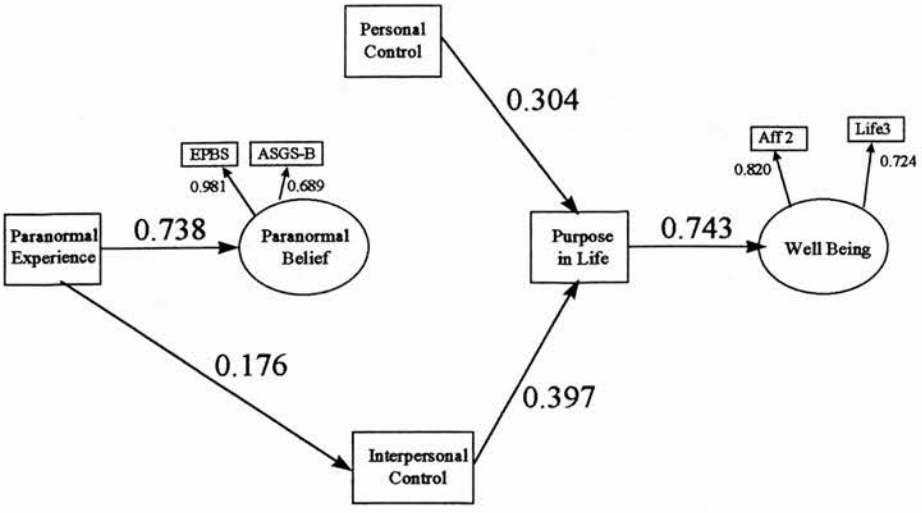

Figure 4. The respecified personal consequences model (MK II) with path coefficients.

the Wald test were implementable the same could not be said for the Lagrange test.

The initial respecified model, with non-significant paths removed (save for the error term for the EPBS), was calculated but a feature of the respecification was a non-significant Wald test result for the link between paranormal belief and personal control. This path was removed from the model, and the model once again recalculated. The results for this second respecification are depicted in Figure 4 and the second column of table 1.

As we can see, the respecified model fits somewhat equivocally, with a significant chi-square goodness of fit statistic, but a reasonable non-normed fit statistic. Furthermore, looking at Figure 4 we can see that the solution offered

is far from acceptable in terms of furthering our understanding of the influence of paranormal belief and experience upon well being and purpose in life. In short, paranormal experience is the only relevant variable to influence control, and it only influences interpersonal control. If this model is to be believed then only paranormal experience has any consequence for purpose in life and well being. This is, frankly, hard to believe from a theoretical perspective, and I am sorely inclined just to ignore this whole solution in favour of the original one. That would mean accepting, at least provisionally, that the influence of paranormal experience and belief upon well being and purpose in life is weak and indirect, but this seems to me to be more realistic than the alternative. I now turn to a discussion of these results.

Discussion

In contrast to previous modelling attempts, the models outlined in this chapter have met with many problems. Though there seems to be some theoretical justification for the notion that paranormal beliefs might enhance coping with uncontrollable life events, the analysis has failed miserably to find confirming evidence for causal models implementing this hypothesis. A model with equivocal fit suggested on the basis of a specification search showed absolutely no role for paranormal belief in influencing purpose in life or well being, directly or otherwise. I am sufficiently perturbed by this as to wonder whether some confounding factor, not included in the model, might be preventing a clearer picture from emerging. One possibility is immediately apparent - that there might be sex differences in the relationship between paranormal factors and well being and purpose in life. If this was the case, and if these sex differences lay in opposite directions, then this may well explain some of the problem in testing the proposed model. The solution would be to analyse the personal consequences model separately for males and females. In part 2 of this chapter I undertake this task, in a purely exploratory way.

TABLE 2A. CORRELATIONS BETWEEN PES AND KEY VARIABLES	Overall Correlations PES (N=129)	Correlation for men PES (N=45)	Correlation for women PES (N=71)
Life 3	0.020	-0.090	0.116
Well Being	-0.031	-0.249	0.101
Personal Control	0.005	-0.158	0.165
Interpersonal Control	0.176	-0.009	0.300
Purpose In Life	0.066	-0.015	0.203

Part 2

Without further ado let us move straight to tables 2a and 2b in which we can see the pearson correlations between paranormal belief, experience, well being and purpose in life for the whole sample, for men, and for women. It is clear from these tables that men and women do express different relationships between paranormal experience, paranormal belief, well being and purpose in life. For men, the correlation between the paranormal variables and the consequence variables is weakly negative. For women, the correlation between the paranormal variables and the consequence variables is weakly positive. The consistency in these correlation differences is striking, and some of them are significantly different from each other on a one-tailed Fisher z-test (Fisher, 1921). Specifically, the Fisher z's for the sex difference in correlations between life 3, well being and paranormal belief scale are significant, as are the Fisher z's for the differences in correlations for well Being, personal control and paranormal experience. Also, the correlations, irrespective of sex, tend to be stronger between the key variables and paranormal experience, than for paranormal belief. Lastly, the overall correlations, irrespective of sex, artificially show no correlation between either paranormal experience and the key variables, or paranormal belief (except for the correlation between paranormal experience and interpersonal control which came through as the only link in the personal consequences model MK II). Clearly, one would do well to analyse the original personal consequences model separately for both sexes. This having been done the results are depicted in table 3 and Figures 5 and 6.

Firstly, no special problems were encountered in testing either the males or females PCM Mk I models (with TOLERANCE set to 0.001). For the test of the male data five parameters were found to be non-significant in the initial analysis; the parameters from belief to personal and interpersonal control, experience to personal and interpersonal control, and interpersonal control to purpose in life. No parameters were suggested for addition according to a Lagrange Multiplier test. In the initial analysis for the female data three parameters were found to be non-significant; the parameters from experience to personal control, and belief to personal and interpersonal control. Table 3 shows the respecified models' fit statistics for the male and female data with these non-significant parameters removed. From table 3 we can see that neither the male or female PCM models give a significant degree of misfit. Given that the PCM seems to fit adequately for both sexes, we need to turn to Figures 5 and 6 to see whether this fit is achieved with different path coefficients.

TABLE 2B. CORRELATIONS BETWEEN EPBS AND KEY VARIABLES	Overall Correlations EPBS (N=129)	Correlation for men EPBS (N=45)	Correlation for women EPBS (N=71)
Life 3	0.034	-0.157	0.169
Well Being	-0.000	-0.301	0.198
Personal Control	-0.061	-0.126	0.058
Interpersonal Control	0.076	-0.048	0.178
Purpose In Life	-0.017	-0.115	0.111

TABLE 3. MODEL FIT STATISTICS	Final Personal Consequences Model Males (N=45) (Fig. 5)	Final Personal Consequences Model Females (N=71) (Fig. 6)
Average off-diagonal absolute standardised residuals	0.0994	0.0914
Independence Model Chi-Square *(df)*	136.139 *(28)*	240.487 *(28)*
Independence AIC	80.139	184.487
Model AIC	-20.984	-9.62603
Chi-Square *(df)*	21.016 *(21)*	28.374 *(19)*
Chi-Square p-value	0.45797	0.07649
Normal theory RLS Chi-Square	17.885	25.340
Bentler-Bonett Non-Normed Fit Index	1.000	0.935

Comparing the two Figures one can see that radical differences in the path coefficients occur in a number of paths. Thus, the path coefficient from experience to interpersonal control is moderately positive for females (0.300) but non-existent for the males. For the effects of the control variables upon purpose in life some particularly interesting differences emerge across sexes. For females and males, personal control is moderately positively related to purpose in life (0.282 and 0.345 respectively). For females, interpersonal control is positively and (moderately) strongly linked with purpose in life (0.529), whilst in males this path is non-existent again. All other aspects of the structural model for males and females have very similar path coefficients. These results are of some significance, and so I shall turn to their discussion.

Part 2 - Discussion

What can we make of the results of this analysis? It seems to me that the results, though exploratory, raise some interesting possibilities. So, for the men we seem to have found something rather drastic - neither paranormal belief or experience are significantly related to issues of personal or interpersonal control, or purpose in life. Some findings of psychological interest may result from the finding that only personal control issues relate to purpose in life for

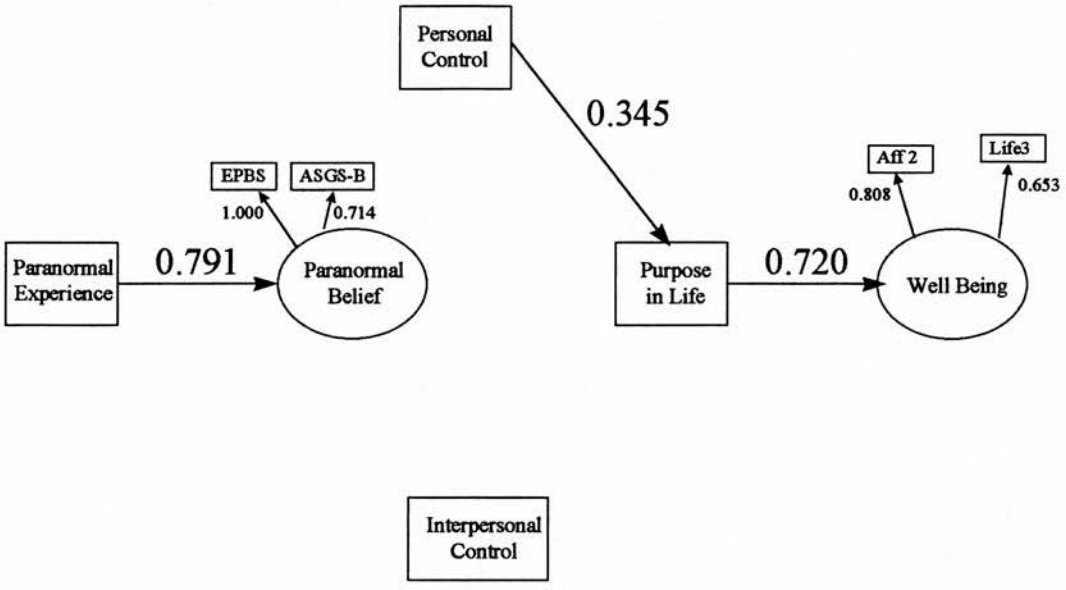

Figure 5. Personal Consequences Model for Males (N=45) with path coefficients.

men, with interpersonal control having no bearing upon purpose in life. For men it seems as though being the master of one's own fate is what influences the sense of purpose in life, whilst having control of one's relations to others is just not relevant. I do not know if this result is present in the literature on purpose in life, but if genuine then I think it should be. As for women, they

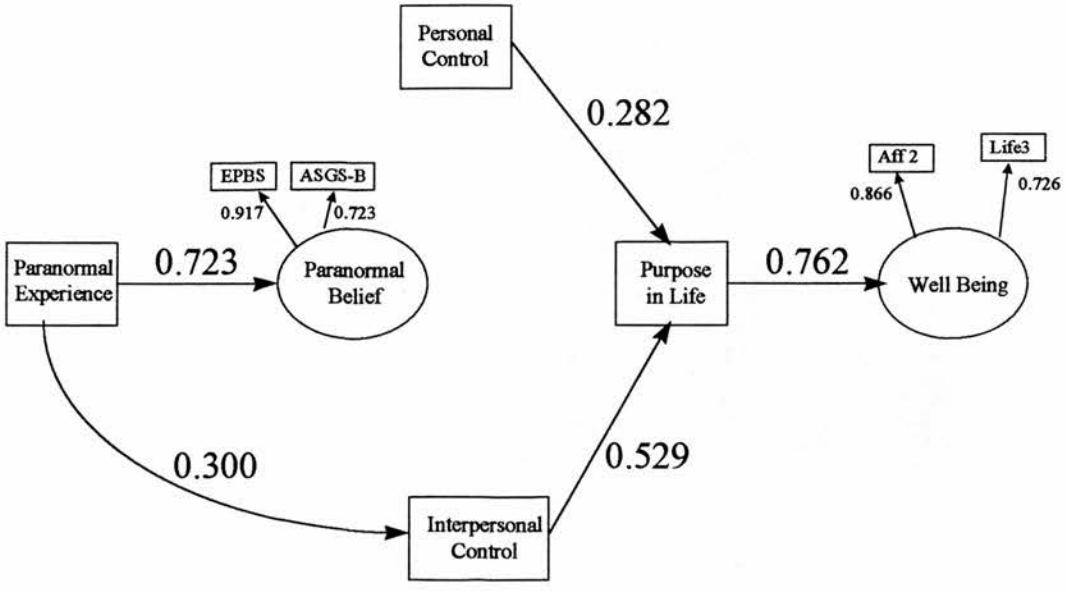

Figure 6. Personal Consequences Model for Females (N=71) with path coefficients.

do display a reasonable relationship between paranormal experience and a sense of interpersonal control. This may perhaps be related to a use of 'women's intuition' in the intuiting of problems or prospects in interpersonal relationships. This is of course pure speculation, but interesting nonetheless. On the psychological front women show reasonable positive relationships between personal and interpersonal control and a sense of purpose in life. For women, in contrast to men, it is interpersonal control which is moderately related to enhanced purpose in life, suggesting that for them purpose in life is partly gained through keeping control over their relationships with others.

Lastly, as we saw from table 2a and 2b men generally show weak negative relationships between the paranormal and control/life meaning/well being variables, whilst women show stronger positive relations. This is an intriguing result, which is at one time surprising and obvious. We know from surveys that women and men differ in terms of their strength of paranormal (psi) beliefs, with women displaying stronger paranormal belief. It is then perhaps only superficially surprising that men and women should display differing relationships with paranormal belief and locus of control. Nevertheless, this discovery considerably modifies, and warrants caution against, the interpretation of previous findings relating paranormal belief to locus of control type measures. Whilst women show positive and moderate strength relationships between paranormal experience and personal and interpersonal control, mean show either weak (interpersonal) or weakly negative (personal) strength relationships with paranormal experience. Even more interesting still, is the discovery that paranormal belief plays a very small role in the determination of personal control or interpersonal control. Once again we see that paranormal experience is the more interesting variable, and that previous studies influenced by a more sceptical orientation to ignore experience have done so at their cost. Nevertheless, how should we explain these sex differences? What needs to be explained is how women seem to benefit from enhanced personal control and interpersonal control when they claim to experience paranormal phenomena, whilst men barely gain any benefit, and indeed are more likely to perceive less control over interpersonal matters when claiming high paranormal experience. In all likelihood a whole Ph.D. could be devoted to an examination of these sex differences, and so what I offer here will perhaps only scratch the surface. It seems one possibility can be drawn from the existing literature. Thus, we know that women claim higher levels of paranormal belief than men (Emmons and Sobal, 1981; Thalbourne, 1981; Tobacyk and Milford, 1981; though Thalbourne and Delin claim their ASGS is indifferent to the sex of the participant). Certainly on Tobacyk's Revised Paranormal Belief Scale there is a tendency for women to score higher than

men on most of the seven subscales except for extra-ordinary lifeforms. It is interesting that ELF is the only scale on which men score more than women, indicating (potentially) that men believe more in the aliens and their spacecraft than do women.

What is interesting about the ELF scale is that it is starkly materialistic in comparison to the other scales. Perhaps then men prefer to believe in things they can see with their real eyes, whereas women gain more from imaginative type paranormal beliefs (ones that depend upon inner space as opposed to outer space). Such belief distinctions may stem from subtle contrasts in cognitive style between men and women in their imaginative capacity. Thus, when women experience an ostensible paranormal phenomena of the mentalistic kind they are able to accommodate it within their conceptual and cognitive system, whilst men are less able to do this, being more driven by propositionalistic styles of thinking. Incapable of integrating such mental intrusions as paranormal phenomena consists of the man may choose to ignore them or may even become concerned at their pathological implications. A vision of impending mental deterioration, and the self doubt this brings, may unless resolved have negative consequences for personal and interpersonal control such as have been seen weakly in our data. The argument offered here goes well beyond any of the present data, and one need not be bound to its essentially cognitive or constitutional premises to explain such sex differences. Social or interpersonal accounts in terms of socially developed coping styles may work just as well, and may even offer more fertile soil for future investigation. Needless to say I think the discovery made here points to a potentially rich avenue for examination.

We have now covered both the head and tail of Irwin's original model of the development of paranormal belief and experience. Interesting discoveries have been made, some substantial and some less so, and there is I feel good reason for thinking that we now have an augmented (I shall leave it to my peers to judge whether improved) understanding of the potential causes and

consequences of paranormal belief and experience. However, it has become clear in the course of this research that paranormal experience is a key variable in all the models developed, and that paranormal belief appears very much more consequential than causal. Up to this point I have been implicitly using the term paranormal experience to mean 'ostensible paranormal experience', but I wish in the next, penultimate chapter to tackle the prefix 'ostensible' by examining more generally the question of whether paranormal belief can be related to laboratory based studies of paranormal experience. A whole host of studies have been conducted since the early 1940's examining the so-called sheep-goat effect in force-choice ESP tests. These studies claim to provide support for the notion that belief may assist in the ability to score well in ESP tests, which if true would provide an interesting reversal of the causal trend developed in this thesis (from experience to belief), not only because it would suggest that belief could sometimes cause paranormal experience, but that it could sometimes causes *real* paranormal experience (albeit of the limited and largely dispositional kind found in forced choice tests). Let us then don our intellectual Wellingtons and Barbour and step in to the farmyard of the parapsychologists' laboratory.

Chapter 10

GATHERING IN THE SHEEP AND GOATS...A META-ANALYSIS OF FORCED CHOICE SHEEP-GOAT ESP EXPERIMENTS, 1947-1993.[1]

For the most part this thesis has pursued a course of indifference with respect to the question of whether paranormal phenomena really exist. Indeed, it is commonly presented as a benefit of research in to paranormal belief that it does not necessitate any active interest in the psi hypothesis. Because of this the area of research in to paranormal belief is the most academically correct of all areas of research in to the paranormal. Despite its pretensions to academic correctness however, research in to paranormal belief that does not stop to ask the question of whether or not paranormal phenomena may be implicated in any way runs the risk of intellectual dishonesty. At this juncture then, just prior to drawing the curtains on this thesis, I want to stop and permit myself the freedom to consider the possibility that the people whose opinions and experiences I have surveyed are not all deluded, stupid, gullible etc. Of course, we can never go back to the circumstances of their individual experiences and adjudicate their paranormality. Nevertheless, the broad thrust of their experiences would suffer somewhat were parapsychologists unable to find any valid evidence for paranormal phenomena. In this chapter then I take it upon myself to examine, in depth, one particularly pertinent area of research, namely research into the so-called sheep-goat effect. In my thesis I have been principally concerned with the causes and consequences of belief in the paranormal and paranormal experience, so what better way to fortify this research than to examine that large(ish) body of experimental studies that have sought, prima facie, to examine the effect of belief in ESP upon one's success

[1] I wish to thank Jessica Utts, Professor of Statistics at UC Davis for statistical help with this meta-analysis.

in experimental tests of ESP. In choosing to examine this area of research I could fulfil two tasks - to examine an area of pertinence to this thesis, and to make a genuine summative contribution to parapsychological knowledge. The technique of meta-analysis has of late become something of a cause celebre of the parapsychologist (Broughton goes so far as to call it a controversy buster!), and it has now been employed to positive effect in a number of important areas of parapsychological research (Utts, 1991). But one area of research that had not been meta-analysed at the start of my research was the sheep-goat effect. And so in this chapter I detail my own attempt to meta-analyse this area of research. As it happens, this enterprise assumes some significance in the context of our wider model building attempt. As I suspected, and as Peter Bentler was able to confirm, one can fix any substantively obtained correlation coefficient in to a path analysis, whether that coefficient comes from data obtained in that path analysis or not. And so, on the face of it, the effect size measure we obtain for the sheep-goat effect has the potential to clear up, for sure, at least one interesting part of the wider model - to what extent can paranormal belief facilitate paranormal experience.

One of the reasons that people who are sceptical of the paranormal often give for their scepticism is that they have never experienced anything that they would deem remotely psychic. Schmeidler (1943) proposed that at least one reason for the sceptics' lack of apparent psi experiences may be that sceptics tend to subconsciously avoid them. Moreover, using forced-choice card guessing designs, Schmeidler (1943; Schmeidler and McConnell, 1958) found that sceptics of the paranormal consistently scored below the mean chance expectation for such experiments. Schmeidler termed those who professed scepticism the *goats* and those who professed belief in the paranormal- *sheep*.

Subsequent research into the so-called *sheep-goat effect* (from now on abbreviated to SGE) by independent investigators confirmed the finding that sheep tended to score above chance on ESP tests, whilst goats tended to score

below. The consistency of this finding lead to a major review of SGE studies in the early seventies, with Palmer (1971) ending his survey by concluding that:

> *"...the data presently available support the hypothesis of a genuine SGE, although the relationship is very slight and difficult to display with small samples. It also is likely to be quite sensitive to the experimental situation...The SGE may indeed prove to be considerably stronger once a more reliable measure of ESP is found." (p.405)*

In this chapter I present the results of a meta-analysis of 73 studies that looked at the influence of belief in ESP on scoring in ESP tests, published between 1947 and 1993. Five major questions are asked: (1) Is there an overall SGE? (2) What is the effect size of the overall effect? (3) Is the display of a SGE related to methodological factors that might suggest a non-paranormal explanation? (4) Does the SGE vary meaningfully with potential moderating variables (e.g. experimental setting, sample size, knowledge of results)? (5) Does the SGE depend upon the type of belief measure used?

The last question is of particular importance given the wide variety of measures of belief in ESP (and also the paranormal in general). Furthermore, at least one study (Palmer, 1972b) has shown that a consistent SGE is likely only when a particular type of question is used, namely Schmeidler's (1943) original question- 'Do you believe it is possible that ESP can be shown under the conditions of this experiment?' (This type of question will be referred to as the

Criterion 1 question, after Palmer, 1971). If we can't separate out the lamb from the mutton how are we going to separate out the sheep from the goats?

Study Search Methods

Source of Studies

Studies were taken from any English language source that addressed the relation of subjects attitude to the paranormal and its influence on ESP scoring. This represents an 'all in' selection policy to try and get as many studies in the analysis as possible. In matter of fact, most studies were from the standard parapsychological journals; *Journal of Parapsychology, Journal* (and *Proceedings*) *of the American Society for Psychical Research, European Journal of Parapsychology* and abstracts of peer reviewed papers presented at the Parapsychological Association meetings published in *Research in Parapsychology.*

Outcome Measures

The first thing that must be noted is that some authors looked at more than one definition of sheep-goat. Palmer (1971; 1972a) identified at least four classes of sheep/goat and some authors employ more than one of them in their analyses. There is therefore a problem of multiple dependent variables, all potentially correlated. The approach adopted here is that favoured by Rosenthal (1984) and simply consists of taking a single measure of the dependent variable wherever possible. In keeping with this approach, I have specifically employed the *method of mean result* (discussed in Rosenthal, 1984). Using this method all measures of outcome (zs and effect sizes) are totalled and divided by the number of such outcome measures to give a mean z and effect

size. However, at the request of Schmeidler (personal communication, 1993) if a study used Schmeidler's original Criterion 1 question amongst other criteria then ONLY Schmeidler's criterion was used in the analysis. Thus this meta-analysis will directly assess the utility of Schmeidler's criterion for separating sheep and goats. Furthermore, it has been common for some researchers to use a three way split of belief into sheep-undecideds[2]-goats. In the Palmer (1971; 1972a) review these agnostic subjects scores were split evenly between sheep and goats. Thus half the hits of undecideds went to sheep and half to goats. This seems a half-baked technique to use to get round the problem of a three way split of belief. The present author feels that, given the fact that we are looking for a *sheep-goat* effect, it is better just to ignore the data from undecided subjects. What is more, there is evidence that undecideds score higher than sheep on the average (Bevan, 1947; Casper, 1957; Nash and Nash, 1967). At least with the approach adopted by this author we can be sure that we are looking at the scoring differences between those who believe and disbelieve in ESP.

Significance levels. Three basic procedures were employed to assess study significance;

1. *Full data given-* where the study gave full data for trials, hits, and probability of a hit for sheep and goats then significance level was determined by calculating the *z-score of the difference in two proportions* using a formula (number 10) given in Edge *et al.* (1986, pp. 154). These calculations were done on a computer.

2. *Other stats given-* where trials, hits and probability of a hit were not given, but some other statistic was used instead (e.g. Pearson's *r*) then if

[2]Also known as 'conflicts'.

possible a transform to z was calculated and used. In most cases this consisted of a transform of r to z using;

$$z = r\sqrt{N}$$

3. No data given- where no data were given by the author the outcome of the study was always described by that author as 'not significant'. Quite simply, these were given z = 0. Eight such studies were found.

Effect size. Once a single significance level had been determined, the effect sizes were calculated. As in Honorton and Ferrari's (1989) meta-analysis of precognition studies a trial based estimate of effect size was used. However, as sheep-goat studies look at mean differences between two groups a modified effect size estimate was calculated as;

$$r = z\sqrt{\frac{1}{N_s} + \frac{1}{N_g}}$$

where N_s and N_g are the number of trials for sheep and goats respectively. This was the procedure used when a study gave full data (trials, hits, probability of a hit). Calculations of effect size on full data were done on computer. Two other procedures were also used to estimate effect size;

1. *Other statistics given-* where some other statistic was given, then that statistic was transformed to a z (using the procedure described for assessing significance given above) and then that z was used to calculate an effect size as for full data. This of course presumed that one could infer accurately the number of trials given to sheep and goats. If this could not be done then the second procedure was used.

2. *No data given-* in all cases where no data were given the author merely describes the SGE as 'non-significant'. The effect size given for such studies is zero.

Overview of Meta-analysed Studies

The author found 73 studies published over a 46 year period from 1947 to 1993. The database consists of over 685,000 forced choice ESP trials with more than 4500 subjects. Study sample sizes range from 140 to 50000 trials (mean = 10539.8; median = 5750). The total number of subjects per study ranges from 9 to 399. Most studies used students. Methods of assessing sheep-goat status varied widely, and the following common measurement questionnaires were included separately for analysis; Schmeidler's *Criterion 1* (Schmeidler, 1943), Thalbourne's *Australian Sheep-Goat Scale* (Thalbourne and Haraldsson, 1980), Van de Castle's *Incompleted Sentence's Questionnaire (ISQ)* (Van de Castle and White, 1955) and Bhadra's *Sheep-Goat Questionnaire (BSGQ)*, (Bhadra, 1966). All other means of assessing sheep-goat status were included in the category 'others'.

Table 1.
Overall Evidence for the SGE

	Mean	S.D.
z-score (N=73)	0.956	1.365
Effect Size *r* (N=73)	0.029	0.045

Combined $z = 8.17$, $p = 1.33*10^{-16}$

"Fail-safe N"= 1726
t(ES) = 5.54, 72 *df*, *p* = .0000002
95% Lower confidence estimate = .019

Combined Evidence for the SGE.

Quite simply, there is a significant overall SGE. From Table 1 (above) it is clear that the overall results are highly significant[3].

Eighteen studies (24%) show significant sheep-goat effects at the 5% level. Z scores correlate positively and almost significantly with the square root of the sample size $r(65) = .218$, $r2 = .048$, *p(1-tailed)* = .081. Indeed, the mean sample size for significant studies (mean = 13742.5) is 47% larger than the mean sample size for non-significant studies (mean = 9313.34). This difference is almost significant; $t_{diff}(63) = 1.379$, *p(1-tailed)* = .0864.

[3] Statistical analyses were performed on Statview on an Apple Mac. All *p*-values are two-tailed unless otherwise stated. Combined *z*s are based on Stouffer's method (Rosenthal, 1984).

Effect Size By Principal Investigator

When the first author of a study is used to assess effect size instead of the study itself a similar result emerges. There are 37 such first authors. The mean effect size for investigators is .026 (S.D. = .043). Even when investigators contributing more than three studies are left out (leaving 35 investigators) then, mean effect size is still .026 (S.D. = .045). Along with previous meta-analytic demonstrations of cross-investigator replicability (e.g. Honorton and Ferrari ,1989) these data provide strong evidence against the view that the results of parapsychological experiments are attributable to a few 'lucky' major contributors (Akers, 1987).

The Filedrawer Problem

The practical, intellectual and even emotional investment in a piece of research is often very great. The need to get results is therefore considerable, and the disappointment that null or opposite results give rise to is understandable. However, there may well be a temptation under such circumstances to pass off the results obtained as just 'bad luck' or seek ad hoc methodological errors to justify ignoring the outcome. In such cases the researcher may, so it is supposed, throw their paper into the bottom of their filedrawers along with all the other studies that didn't 'get results'. Hence, the failure to publish null results is potentially very important when summarising the outcomes of published studies in any area of behavioural research. This so called 'Filedrawer problem' (Rosenthal, 1984) is very pertinent for parapsychological research. Indeed, in 1975 the Parapsychological Association adopted an official policy urging researchers to publish null results as well as positive.

Table 1 gives Rosenthal's (1984) "fail-safe N" statistic in an attempt to estimate

the number of unreported studies with null results (i.e. average zs of zero) that would have to be sitting in researcher's filedrawers for the reported database effect to be reduced to nonsignificance. In terms of the SGE database, 23 unreported studies averaging null results need to be 'out there' for every reported study. Furthermore, using Dawes *et al.* (1984) truncated normal curve analysis it is possible to provide a complementary measure of the filedrawer severity. Essentially, the assumption is that due to publication and reporting bias the reported studies are only those which have been randomly sampled (assuming a normal distribution with mean = 0 and standard deviation = 1) from a critical level of significance greater than z = 1.65. On this assumption the mean expected z is 2.06, and the variance above the truncated normal of z = 1.65 is .14. For the sheep-goat database there are 18 studies with zs >1.65. The average z of these eighteen studies is 2.68, not 2.06. A test z comparing the actual mean of 2.68 with the expected mean of 2.06 gives a z of 7.03 (.62 divided by (.14/18)1/2). p for this z is .00000110 based on a t(17) distribution.

Given that the actual mean z of significant SGE studies is significantly larger than that expected if such studies were randomly sampled above z = 1.65 on 'no effect' distributions then we can safely assume the existence of some real SGE independent of publication or reporting biases. What is more the "fail-safe N" for the combined SGE is sufficiently large to support the conclusion that even if every professional parapsychologist reported one null SGE study every year for the next seven years the result obtained in this analysis would only then become 'just significant'. Whatever explains the SGE, it is not to be found in the contents of parapsychologists' filedrawers.

Outlier Reduction

The database under scrutiny does not show the serious problem of outlier studies that previous databases have (e.g. Honorton and Ferrari, 1989). There is

however, still over a 9 sigma spread between lowest and highest z-scores. Performing a 10% trim on the data (Barnett and Lewis, 1978) gives 59 remaining studies. The combined stouffer Z for the trimmed sample is 6.55, p = $4.73*10^{-12}$. The trimmed mean effect size is .025, S.D. = .003. The lower 95% confidence estimate for this sample is still .019. Kurtosis (*g2*) is reduced from over three to -1.31. The filedrawer estimate for this trimmed sample is 872 studies (14 null for every known one). Where there is a serious problem with outliers it is sensible to perform subsequent analysis on the trimmed sample. In the case of the sheep-goat database there does not seem to be a severe problem with outlying studies. It therefore seems appropriate to perform study quality analyses upon the original sample of studies.

Study Quality

One of the main criticisms offered by critics of parapsychology is that the significant effects depend upon the worst controlled studies. If studies are protected against sensory leakage, given adequate randomisation and have duplicate or automatic recording and checking of guesses then no psi will occur. Subject fraud and experimenter error will have been minimised.

In all types of meta-analysis the studies under scrutiny will vary widely in quality- along a number of dimensions. How to deal with this variation is a point of some argument amongst meta-analysts. In this meta-analysis the author has chosen to look at all studies on the SGE (in the forced choice domain). Whilst discarding studies that are not up to scratch is the crude equivalent of assigning them a quality weighting of zero it should be remembered that if a study is physically thrown out of the analysis all of its other potentially useful information is gone. This author has chosen to analyse studies with weights of zero all the same, should they arise.

Study Quality Criteria

As for Honorton and Ferrari's (1989) meta-analysis, research quality was assessed in terms of a number of procedural descriptions, giving details on methodology and statistics. Essentially, one point is given for each of the following criteria:

Specification of sample size. If the investigator explicitly pre-specifies the sample size then credit is given. If the study uses group testing- where it is not possible to pre-specify the sample size precisely, then credit is also given. Credit is withheld for studies in which the sample size is not planned or not mentioned.

Preplanned analysis. If the method of statistical analysis, including the outcome (dependent) variable is preplanned then credit is given. If neither the experimental design or analysis are mentioned in respect of the dependent variable of concern then no credit is given.

Randomisation method. Credit is given if the study was adequately randomised. Adequate randomisation techniques are; random number generators, random number tables, and mechanical shufflers. Hand-shuffling, drawing lots and die casting do NOT constitute adequate randomisation and are not given credit.

Controls. If studies report randomness control checks (or control series) or empirical cross-check controls then credit is given.

Recording. If target sequences and subject responses are automatically recorded, or recorded in duplicate (e.g. via carbon copy) a point is awarded.

Checking. If checking of target-response matches is done automatically or in duplicate (e.g. by two independent checkers) then credit is given.

Controls against sensory leakage. Where studies have no sender (i.e. clairvoyance studies) then one point (and only one) is given if any or all of the following are evident; 1) Opaque container for targets/RAM memory stored targets. 2) Subject screened and blind to experimenter. 3) Subject is physically remote. *In the case of precognition a point is always awarded as no sensory leakage is possible.* In the case of studies with a sender then one point (and only one) is awarded if any or all of the following are evident; 1) Different rooms separate sender/receiver. 2) Room shielding documented. 3) 1-way communications from receiver documented. 4) Sender/Receiver monitored vigilantly.

Study Quality Analysis

A total of seven quality points are available. The range of points awarded was 1-6 (mean = 3.164, S.D. = 1.247). There is no significant relationship between study quality and effect size. The quality weighted stouffer $z = 7.97$, $p = 5.07*10^{-16}$. The weighted effect size is 0.022.

Quality Extremes

When studies are split into quality extremes is there a tendency for very poor studies to get larger effect sizes than extremely good studies? 10 good studies (quality ratings of 5-6) had a mean effect size of .018 (S.D. = .063) and 23 poor studies (quality ratings of 1-2) had a mean effect size of .032 (S.D. = .055);

$t_{diff}(31) = -.631, p = .5325$. Thus, whilst the tendency for poorer studies to get larger effects is shown, the difference is not significant.

Quality Variation in Publication Sources

Quality of studies does vary depending upon the journal the study was published in; a Kruskall-Wallis One-Way ANOVA corrected for ties gives H(6) = 23.27, p = .0007. This is due largely to the poorer quality of studies published in *Research in Parapsychology*. However, this result will depend to some extent on the fact that RIP articles tend to be abstracted and don't give as much information as fuller journal articles. Despite this, there is no significant difference between effect size and publication source on a One-Way ANOVA; $F(6, 72) = .255, p = .9554$.

Study Quality in Relation to Year of Publication

Study quality has not increased over the years; $r(72) = .017, p = .8889$. Effect size has, however, slightly declined over the years; $r(72) = -.083, t(72) = .705, p = .4833$.

Table 2.
The Effect of Belief Scale on Sheep-Goat Effect Sizes

Belief Measure	N(Studies)	Mean(ES)	S.D.
Bhadra (BSGQ)	5	.027	.043
Australian	6	.032	.063
ISQ	5	.011	.079
Criterion 1	30	.032	.036
Others	27	.029	.046

Overall main effect; $F(4, 72) = 0.239$, $p = .915$

The results of the study quality analysis are clear and confirm the meta-analyses of previous authors. The sheep-goat effect is not related to the quality of the study. Yet again the view that only the worst controlled studies would display ESP effects is shown to be wrong.

Moderating Variables

Comparisons of Belief Assessment Efficacy

The question addressed here is one of considerable importance for future research in this area. Does the size of the SGE depend upon how one measures subjects' belief in the paranormal? Palmer (1972b) suggested that only Schmeidler's (1943) Criterion 1 question would give consistent results. Furthermore, Broughton (1991) has shown that the number of subjects classified as sheep and goats depends upon the type of question they are asked, and varies widely within subjects and across questions. From my own

experience, it is clear that a sheep on one scale may be a goat on another. Despite this potential source of confusion, the results are surprisingly coherent and consistent.

It is clear from Table 2 that there is no overall relationship between type of belief measure on effect size. Only the ISQ shows any likelihood of being a poor measure of the SGE, yet it is not significantly smaller than any other measure on a Scheffe F-test. More importantly, there is no apparent superiority for Criterion 1 over other sheep-goat scales, although it does give the joint highest mean effect size, with the Australian Sheep-goat scale (Thalbourne and Haraldsson, 1980). Some measures of belief in the paranormal are to be noted for their absence. The Paranormal Belief Scale (Tobacyk and Milford, 1983) and the Belief in the Paranormal Scale (Jones *et al.* 1977) are both potentially useful measures of sheep-goat belief, and they have been profitably used in psychologically oriented studies of belief in the paranormal (see Irwin, 1993, for a review). It appears that these measures have failed to cross over to parapsychology. Whether this is a good or bad thing is something for later discussion.

The results of this section are clear. There is no single best way of separating sheep from goats. All measures used to date give approximately the same answer. In many respects this is good news- the SGE seems quite robust.

Other Moderating Variables

In their meta-analysis of the forced-choice precognition literature, Honorton and Ferrari (1989) found two variables that systematically covaried with effect size that are also of potential importance to the SGE database; 1) Individual vs. Group testing, 2) Feedback of Results.

Table 3.
Feedback Received By Subjects In Sheep-Goat Studies

	Feedback Type			
	None	**Session**	**Trial**	**Run-Score**
N(Studies)	1	6	2	22
Mean (*ES*)	-0.02	0.033	-0.035	0.029
SD (*ES*)	-	0.044	0.124	0.032

Overall main effect; $F(2, 27) = 2.268$, $p = 0.123$
(Data for no feedback not included in ANOVA)

Furthermore, insofar as SGE studies are addressing personality factors we might reasonably expect variables that influence these personality factors to also affect the SGE. Most pertinent is a finding by Honorton and Bem (1991) that an extrovert-introvert correlation with ESP scoring was only demonstrated in forced-choice data when the subjects were given the personality measures after they were tested for ESP. This suggests that the subjects' knowledge of their ESP score may influence their level of reported extroversion post-experimentally, and also raises the possibility that the experimenter may influence the subject after finding out their ESP score-eliciting certain personality type responses over others. Therefore, I look at whether the sheep-goat scale was given before or after the ESP test, with the prediction that the SGE will be significantly higher in the 'after' condition than the 'before' condition.

It is predicted that feedback will result in enhanced ESP performance, and that trial by trial feedback will be better than other types of feedback, as per Honorton and Ferrari (1989). As with Honorton and Ferrari it is also predicted that individually tested studies will give larger effect sizes than group studies.

Feedback

Firstly, it should be said that no firm conclusions are possible from the data presented in Table 3. The number of studies explicitly stating that they looked at 'no feedback' and 'trial-by-trial feedback' is insufficient. Therefore, the crucial confirmation of Honorton and Ferrari's previous data (showing superior effect size for trial feedback over run over none) is not feasible.

Individual Versus Group Testing

Honorton and Ferrari (1989) found that only studies testing subjects individually gave significant non-zero effect sizes. They argued that group studies were often conducted 'off the cuff' and with little preparation and presented evidence to suggest that such studies were of significantly lower quality than their individually tested counterparts. The likelihood is that group studies don't invest the time and personal care that individual studies do. However, from Table 4 we can see that in the SGE database neither of these results is found.

The results clearly show no major difference between group and individual studies in terms of effect size and a superiority for group studies in terms of quality (p = 0.1301). But perhaps this makes some sense- sheep-goat distinctions characterise to some extent group differences. Points of disagreement come to the fore when one is matched against a group of direct opponents and given a chance to prove one's viewpoint. Perhaps the group setting is more conducive to the sheep-goat effect, overriding the lack of personal attention given in group studies.

Before Versus After (Belief Measurement)

Honorton, Ferrari and Bem (1991) only found a significant association between

extroversion and psi scoring for forced-choice data when the subjects' personality traits were assessed after the ESP test. Sixty-five studies gave information as to whether the S-G scale was given before or after the ESP test, and the results loosely support the conclusion that psi effects are larger when the S-G scale is given after (N = 15, mean ES = .044, SD = .071) than when subjects are assessed before (N= 50, mean = .028, SD = .037) testing for psi. Although a *t* test does not give a significant result (p = .248), the degree of the difference is nevertheless cause for concern. Perhaps it would be best if all future research into the SGE administered the S-G scales before any ESP tests.

If the results of this section on moderating variables say anything it is that the SGE is surprisingly consistent across a wide range of situations and experimental variations.

Predictive Consistency in the SGE Database

In his major 1971 review of the sheep-goat effect, Palmer concluded his discussion of the magnitude of the sheep-goat effect by predicting the percentage outcome of future studies (on the assumption that the average difference between sheep and goats in terms of run score was 5.10 to 4.95 (or .15). In these terms he concluded that;

1. About 84% of studies would yield SGE's in the predicted direction (S>G).
2. About 16% of studies would yield SGE's that were significant in the right direction.

3. About 1 % of studies would yield SGE's that were significantly reversed.

   Therefore...

4. About 15% of studies would yield SGE's that are reversed but not significantly so.

From Figure 1 we can see that Palmer's (1971) predictions were fairly accurate. A further 25 studies looking at the SGE appeared after Palmer's review. Only studies giving full or translatable statistics are included, unless a study gave the direction of non-significant results. The reduced sample comprises 60 studies. Palmer's predictions are probably a bit out because of his use of the score splitting method for studies using 'undecideds', which tends to underestimate the number of significant SGE's. Nevertheless this meta-analysis provides considerable consistency with an updated database to predictions made from a literature review conducted over twenty years ago. Clearly the SGE is a robust and consistent one.

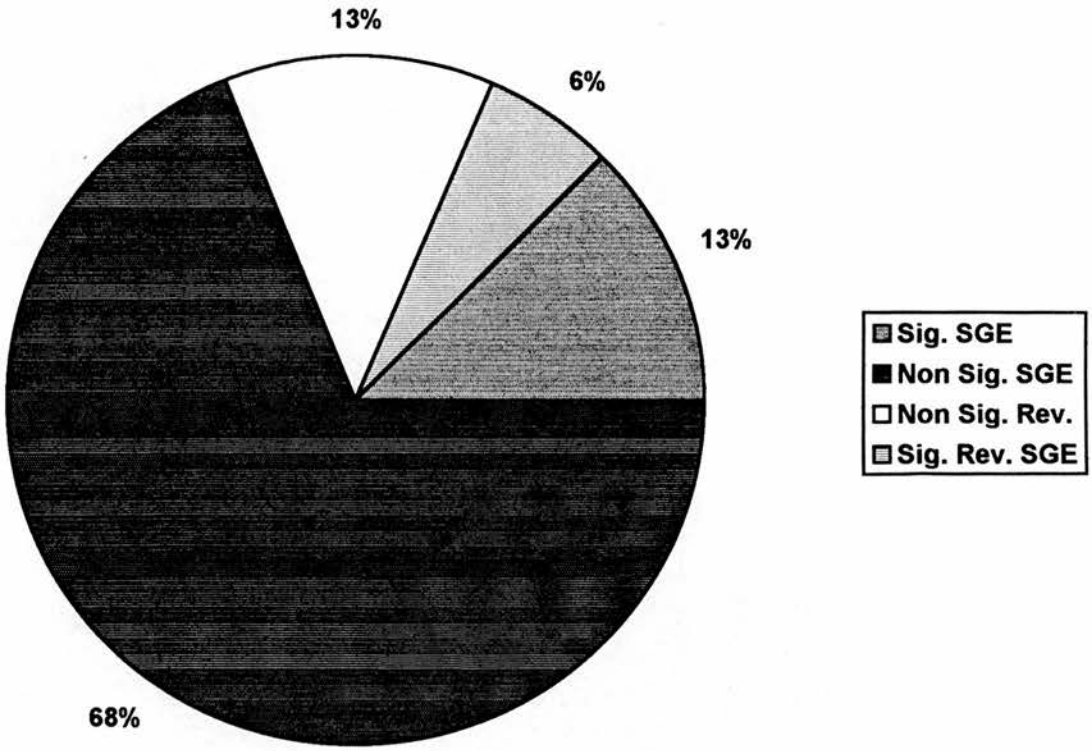

Figure 1.SGE study outcomes for studies conducted after Palmer (1971) (N=16).

Summary and Conclusions

What is reported here provides some considerable agreement with previous meta-analyses. And like those previous meta-analyses the findings indicate the existence of a small, but highly significant effect. The degree of replicability found here both in studies and across investigators is again something that is not new amongst meta-analyses of psi phenomena. No longer can the critic claim that replicability is not to be found in parapsychology- although they may

still staunchly defend the primacy of the $p<.05$.

More specifically, the results of this meta-analysis are quite clear- if you believe in the paranormal *you will score higher on average in forced choice ESP tests than someone who does not.* That the SGE does not depend upon the quality of the research is not surprising in light of the number of previous meta-analyses of psi phenomena that have shown just the same independence from these sources of artefact (see Utts, 1991, for a good review of meta-analyses in this field).

Nevertheless the critic might argue that despite all the significance the actual effect itself (0.029) is too small to be of any real world importance. One would be wrong to suppose that such small effects are of no consequence however. For example, in a report published by the Steering Committee of the Physicians' Health Study Research Group (1988) an experiment looking at the effects of aspirin in preventing heart attacks was terminated after the beneficial effect of aspirin became obvious- aspirin patients clearly had fewer heart attacks than non-aspirin patients (placebo control group). The effect size was 0.03. Another statistic to answer just this type of criticism was devised by Rosenthal and Rubin (1982) and is known as the *Binomial Effect Size Display* (BESD). In the case of a SGE size of 0.029 the BESD shows that 51.5% of sheep would benefit from a gain in score simply as a result of their positive attitude to the paranormal, whereas only 48.5% of goats would show a gain in ESP score in spite of their scepticism. Seen in these terms the weak sheep-goat effect looks far more impressive.

The finding that the SGE is relatively robust across different measures is also cause for some relief. However, this does not mean that researchers should continue with their 'pet' SG scales. What is needed is a good, reliable, accurately validated measure of general belief in the paranormal (with sub-scales for a psi component). Questions should most certainly include the Schmeidler question seen to be joint most successful measure of belief in terms

of getting results. The need for such a universally accepted scale is more important now that psychologists and parapsychologists are looking in to the wider consequences of belief in the paranormal (Irwin, 1993). It is sadly the case that no current measure of belief (either in parapsychology or psychology) is good enough to encompass the requirements for universal acceptance. What is more, there are at least two schools of research on belief in the paranormal. The first school is represented by the research reported here. Researchers in this mould are broadly open to the possibility of genuine paranormal phenomena and report their results in parapsychological journals. On the other hand are those researchers who can broadly be characterised by their interest in what Irwin (1993) titles - *cognitive deficits research.* Such research is carried out by sceptics of paranormal claims, mostly psychologists, and is reported in psychological journals. The essential question for these researchers is 'How can people come to believe in the paranormal in spite of its existence?'. Whilst these two schools of thought often seem incommensurable, in fact a pragmatically minded approach to belief in the paranormal must find both of them of equal utility in guiding our research. It is a shame therefore that there is such little cross-talk between these two schools. There are some instances of research that are clearly guided by ideas from both these schools (e.g. Irwin, 1992; Lawrence, 1991) but for the most part a conceptual chasm exists between the psychology and parapsychology of belief in the paranormal. Whatever one may personally think regarding the existence of psi, there are clearly opportunities being missed here.

At this point I feel it is necessary to issue some caution over the accepted interpretation of the SGE. Typically the SGE is interpreted as an effect of belief upon scoring in ESP tests. However, such confidence in the causal process can only be claimed if the experimenter has control over the independent variable in the experiment and is able to randomly allocate participants to the various levels of the manipulated IV. These conditions do not pertain in the standard SGE study. In the standard study, which accounts

for most of the studies reported in this area, the researcher takes a sample of people, administers some measure of paranormal belief to them, and then splits the sample in to two groups sheep or goat on the basis of a mean or median split. What one has in this kind of study is an ersatz IV 'fixed' at two levels with no random assignment of participants to these levels. If the participants were selected for the study as a whole at random then in reality we have an arbitrary dichotomisation of randomly distributed belief scores being associated with a random distribution of ESP test scores. But it is a simple matter of fact that the dichotomisation of a sample of random scores reduces information in the data drastically and thus reduces sensitivity of parametric statistical tests calculated on these data. With two sets of randomly sampled scores where a natural variation is present over a reasonable range the optimal analysis is a correlational one (i.e. a Pearson product-moment correlation coefficient). Any other analysis (an independent t-test or z-score of the difference in two proportions) is inevitably less powerful. However, this methodological point aside, our discussion reveals a deeper problem with attributing causal priority to belief in such 'experiments'. If all that can be demonstrated is a correlation we have no basis, within the context of the particular SGE study, to suppose that it is belief that is causing ESP test score differences. An alternative conjecture, and one that fits with the broad thrust of this thesis, would have it that ESP test score is an indicator of the latent variable 'general psychic ability' or 'paranormal experience', and as we know psychic ability does very much have a causal role to play in the development of paranormal belief. Thus we have a situation where general paranormal experience is the co-cause of both paranormal belief and ESP test score. Of course, this implies that there may be no actual direct relationship between belief in the paranormal and ESP test score.

However, there are a few studies which do meet requirements for a genuine attempt to examine the influence of belief on ESP test score. Thus, Lovitts (1983) used an experimental design in which a group of believers and

disbelievers were lead to believe either that they were taking part in a test if ESP (the truth) or an examination of the claim that ESP might result from subliminal sensory cues (a plausible decoy condition - in fact the only way success could occur at guessing targets is if ESP existed). Thus, in one condition success would be interpreted by the participant as evidence for ESP, whilst in the other it could be passed off (in the participant's mind) as merely a subliminal cueing effect. Lovitts supposed that sheep and goats in the ESP condition would show the standard SGE, whilst those in the subliminal cueing condition would show a reversed SGE. Goats in the latter condition would be inclined to score well to vindicate their view that ESP was just subliminal cueing (as a result of sensory leakage) whilst sheep would score badly in an attempt to thwart this hypothesis (presumably by showing a defensive reaction to the subliminal cues). Lovitts did indeed find just such a reversed SGE in the subliminal cueing condition, as revealed by a significant cross-over interaction in a 2-way factorial ANOVA. These results, though interesting, must be treated with caution. There have, surprisingly, been few attempts to replicate Lovitts' work. Thus, Lawrence (1992) failed outright to find a reversed sheep-goat effect in a fairly exact replication. More recently Lawrence and Norman (1997) have utilised a less cumbersome auditory cueing decoy (as opposed to Lovitts' tachistoscopic visual cueing decoy) and did find a near-significant cross-over interaction of just the kind found by Lovitts. Clearly more studies need to be conducted with larger samples, but it seems that belief might have motivationally moderated effects upon ESP scoring after all.

So, cautions aside, what about the prospects for future SGE research? In recent years there has been little research looking directly at the SGE. Is it old hat now? I think not. What the research summarised here has done is establish a small effect in a contrived setting- the forced choice paradigm was never known for its ability to induce excitement. More recently developed free-response techniques have been relatively unused as means to examine the SGE

(Stanford, 1984). One reason why people steer clear of the SGE is that they prefer not to have subjects that deflate their psi scores. But such an approach is clearly short sighted. As much useful information can be gained by planned analysis of data from subjects with opposing views on the paranormal as can be gained from looking at the scores of those who fully accept psi in their lives. In fact, if no one bothers to look at the SGE in free response settings then a whole area of important research, adding useful insights into our understanding of psi missing, is missed. Thus at least one area for future SGE research is to look at it in free response settings. And a specific goal of this research might be to analyse goats' mentations to get a grasp of what psi missing in more ecologically valid circumstances actually looks like.

The odd thing is if the results of meta-analyses like this bear fruit and lead to future research that continues to support the psi hypothesis then public and academic acceptance of the idea of psi would result in the extinction of the sheep-goat distinction- everyone would become some sort of sheep. The SGE may yet become another one of the behavioural sciences' celebrated socio-historical phenomena; a quirky aspect of an era in which science was unable to capture adequately the elusive quarry of psi. Belief in the paranormal may well help you score better...

...but can it, ultimately, help save the sheep-goat effect?

Chapter 11

CONCLUSIONS

> *'His utter dependence on circumstances and environment, the manifold distresses and tribulations of his life, surrounded by hostile neighbours, dangerous beasts of prey, and often exposed to the pitiless forces of nature; his keen senses, his cupidity, his uncontrolled emotions - all these things bind him to the physical realities, so that he is in constant danger of adopting a purely materialistic attitude and becoming degenerate. His belief in spirits, or rather, his awareness of the spiritual world, pulls him again and again out of that bondage in which his senses would hold him; it forces on him the certainty of a spiritual reality whose laws he must observe as carefully and as guardedly as the laws of his physical environment.' - Jung*

In this concluding chapter my discussion will fall in to three parts. Firstly, I wish to consider the empirical data accrued in support of the childhood factors model of paranormal belief development. Secondly, I wish to consider the critique of the paranormal belief scale offered in this thesis, and its consequences for future research in this area. Thirdly, I wish to consider the evidence for genuine paranormal experience as taken from my sheep-goat meta-analysis, and from parapsychology in general. Lastly, I want to explore some possibilities for future research before offering my own view of where research in to paranormal belief stands as a result of my work.

Causes and consequences of paranormal belief and experience

The prescience of Jung appears to have been vindicated in this thesis. Paranormal beliefs and experiences do appear to have a functional role to play in human development, and Jung, with his emphasis on spiritual awareness rather than belief, seems to usher in my own finding that it is paranormal experience which 'does the driving' as it were. The essential findings from the modelling research seem to be thus ;

1) Trauma has a role to play in the development of paranormal belief and experience. Not a new finding, but we have extended the result thus;

   i) it would appear that control related trauma is responsible for heightened paranormal experience via fantasy. This discovery is also of potential relevance for psychologists interested in the traumatic causes of fantasy ability (e.g. Lynn and Rhue, 1988; Bryant, 1995). I do not know of any studies in psychology that have looked for or found this relationship.

   ii) bereavement/loss related trauma is directly related to paranormal experience - this route echoes the findings of psychological studies of the 'hallucinations of widowhood' (Rees, 1971).

2) Fantasy clearly and intuitively has a direct influence upon reported level of subjective paranormal experience. This fits with previous studies, and is the key piece of evidence in support of a psychodynamic theory of paranormal belief development. Thus, it suggests that people traumatised may resort to fantasy as an escape from a harsh reality (in keeping with Freud, 1959/1908; and Lynn and Rhue, 1988), and in so doing develop a capacity for absorbed imaginal states of the kind conducive to paranormal experience (in keeping with recent findings by Ellason and Ross, 1997). In psychology, the traumatic causes of fantasy can be seen in an interesting little study of paracosms, the fantasy universes of children, by Cohen and MacKeith (1991). Thus, one woman describing her childhood paracosm 'Branmail' says that constructing a fantasy world enabled her to 'avoid noticing' her problems and to 'withdraw from the real world where I felt I could not influence events'. However, trauma does not always have to be dramatic to cause fantasy. One child in Cohen and MacKeith (1991) developed a toy world because 'It gives me something to think and dream about when I'm bored.' I suspect that isolation plays its direct role on fantasy (as seen in chapter 8) through 'trauma' of this kind.

3) Paranormal belief is mostly consequential. Paranormal belief has played a small role in the models we have developed. In the trauma models paranormal belief has been the end product of a chain of events in which fantasy and paranormal experience have been central. In the personal consequences model where paranormal belief was, a priori, expected to be a cause of heightened interpersonal and personal control we once again found that beliefs played almost no role. The lack of influence for paranormal belief rather stands in stark contrast to the amount of academic work looking solely at paranormal belief. Psychologists in particular, it would appear, have overestimated its importance. Or perhaps I have failed to include variables with which belief might have more influence.

4) Women seem to benefit from paranormal experience, but men don't. This conclusion comes from an examination of the personal consequences model in chapter 9. There we found;

   a) men and women respond differently to paranormal experience in terms of their level of interpersonal and personal control, thus;

   b) women show moderate positive path coefficients for the effects of paranormal experience upon interpersonal control.

   c) men show a non-significant weakly negative effect of paranormal experience on personal control, and a near-zero positive effect of paranormal experience upon interpersonal control.

   d) both sexes show no significant effects of paranormal belief upon both interpersonal and personal control, those weak effects that are present are all negative.

   e) interpersonal and personal control are both positively related to purpose in life, reinforcing the view that those with higher internal locus of control in those spheres will feel a greater sense of life meaning.

There are some genuinely interesting and original contributions within the empirical modelling work reported within the thesis, specifically with respect to the specified trauma model, and the personal consequences model. For the latter, I have uncovered a sex difference which had previously gone unnoticed, and which may have influenced the strength of relationships found between belief in the paranormal and locus of control in earlier studies. However, the discussion of the results so far has presumed their validity. There may well be problems with the studies reported herein that would invalidate or moderate these conclusions. I outline the main problems, as far as I can see them, below;

1) *The samples used for the studies are too small.* There are two ways in which sample size might influence the covariance structure models reported in this thesis;

    i) The studies reported in this thesis depend upon samples of size 80, 58, and 129 respectively. The studies samples sizes are on the small side, but we should remember that large sample studies in covariance structure modelling, by which I mean studies which give results exceptionally close to the asymptotic properties of the typical fit statistics, are around the 5000 mark. With response rates for the third study as they were, plus time constraints, and costs, the samples obtained probably were the best one could practically obtain. Furthermore, the effect of sample size is felt most directly upon the standard errors of statistics, and small standard errors typically effect such things as cross-validation of models, so one more pertinent question would be how consistent the models were. To the extent that the repeated modelling attempts reported here show the same overall pattern then we may feel confident that small to medium sample sizes have not been a problem. I think we have had a reasonable degree of confirmation for the results where replication in other samples has been attempted. However, for the specified trauma model and the personal consequences model, it

Table 1. Study	Sample Size	Number of freely estimated parameters	N/p ratio
Edinburgh student	80	4	20:1
SPR	58	4	14:1
Edinburgh Basic Model	129	4	32:1
Spec. Trauma	129	17	7.58:1
Pers. Cons. Model	129	12	10.75:1

might be best to exercise more inferential caution, though both those models were tested on general population data and with the largest sample size.

ii) Sample size is critical in the sense that there should at least be more people in the study than parameters to be estimated (the so called sample size: parameter ratio). As Bentler (1989) notes, 'An over simplified guideline regarding the trustworthiness of solutions and parameter estimates might be the following. The ratio of sample size to number of free parameters to be estimated may be able to go as low as 5:1 under normal and elliptical theory. Although there is little experience on which to base a recommendation, a ratio of at least 10:1 may be more appropriate for arbitrary distributions. These ratios need to be larger to obtain trustworthy z-tests on the significance of parameters, and still larger to yield correct model evaluation chi-square probabilities.' (p.6).

It might be instructive to examine a table (Table 1.) of the sample size:parameter ratio for each study. At the worst then one should approach the results of the specified trauma model with caution, whilst the tests of the basic Irwin inspired model of childhood factors developed in chapter 6 and replicated in chapter 7 and 8 can be regarded with most confidence. In any case, whilst no self-respecting person would turn away a large N, the sample sizes don't seem so small for the task set. As Kline (1993) notes with respect to sample size in

exploratory factor analyses, 'Below 100 any factors require replication.' (p. 121). We have seen three replications, with reasonably consistent results for the original childhood factors model based on sample size less than 100.

2) *The samples used are unrepresentative.* I shall examine each sample separately for this problem;

   i) The Edinburgh student sample - this sample is certainly not representative of the general population. It is restricted in age, IQ, and socio-economic background for certain, and could well be unrepresentative in other respects undetermined. However, for me the real issue is whether the unrepresentativeness of the sample is in any way relevant to the study hypothesis. I don't actually think, though I grant it is somewhat presumptive, that restrictions of the kind we see in the Edinburgh student sample are likely to influence my data in a way which could negate my conclusions. Thus, let us suppose that this sample is unrepresentative in such a way that it tends to have a restriction in range on some variable that correlates in a straightforward linear way with paranormal belief and experience (childhood trauma might be a highly plausible example). The effect of this kind of restriction in range is usually to reduce the estimated strength of the correlation coefficient. Thus, if anything relationships found between trauma and fantasy and trauma and experience in the Edinburgh student sample should be *larger* in other more representative samples. The model developed herein would only be supported more strongly. The only way in which the correlations in the Edinburgh student study could overestimate the effects modelled is if the real world correlations were quadratic or of similarly non-linear disposition and our sample were restricted in range to one pseudo-linear component of the quadratic curve. But then again we would only be learning one interesting thing in lieu of one even more interesting thing. That would not be tantamount to

finding no correlation, which would only happen by type II error or a true null. Thus, whilst the Edinburgh student sample may be unrepresentative, I don't think it is so in a way which would seriously challenge the results.

ii) The SPR sample - on first appearances this sample is dangerous. It has both a low N and a highly restrictive range in terms of IQ and socio-economic status. Once again the same types of counter-response apply to this criticism as did for the Edinburgh student sample. Thus, the main problem is restriction of range and unless there really is no effect between the variables modelled the SPR study is unlikely to lead to highly inaccurate inferences about the nature of the effects modelled. Indeed, as we saw, the SPR study broadly replicated the Edinburgh student study. However, the Edinburgh citizen study is not so wildly restricted and shares the same questionnaire as the SPR study, so we *can* compare the descriptive statistics for these two, to look for a restriction of range;

Table. 2 Statistics	**SPR Study (N=58)**				**Edinburgh Survey (N=129)**			
	PES	**EPBS**	**STCE**	**CF**	**PES**	**EPBS**	**STCE**	**CF**
Mean	48.0	37.0	41.3	6.9	32.6	26.0	40.2	6.5
Variance (n-1)	325	56	104	10	272	87	53	11
pooled t-test	5.65*	8.18*	0.86	0.72	**	**	**	**
Starred data are p<0.05								
F-test (variances)	1.19	1.55	1.96	1.10	**	**	**	**

i) The data in Table 2. suggest that the only respect in which the SPR study differs from the Edinburgh survey is in the rather uninteresting fact that the SPR members show a higher level of paranormal experience and belief! Certainly Fisher ratios do not lend support to the view that the SPR sample has a restriction of range on any of these key variables. Unfortunately we cannot compare these data with the original Edinburgh student study as that study

used slightly different questions for paranormal experience. However, I am satisfied that, insofar as restrictions in range can be assessed empirically, there are no serious differences in the two studies that might confound the interpretation of models. However, this analysis depends upon the assumption that the Edinburgh survey provides the most representative survey.

ii) The Edinburgh survey's representativeness - the Edinburgh survey was an attempt to survey a fairly representative cross-section of the British population. Certainly Edinburghers may have their own distinct characteristics, but I am not aware of any evidence to suggest the people of Edinburgh are any more likely to be unrepresentative where paranormal issues are concerned than other group of people in Britain (though logically absence of evidence is not evidence for absence). However, the Edinburgh survey sample is unlikely to be fully representative for a number of reasons. Firstly, it isn't so large that all society's extreme cases could be represented. Secondly, I consciously avoided certain selected streets if they were in areas that looked dangerous (e.g. Niddrie). A strange, young, male, English academic may not be every disenfranchised Scot's cup of tea. So certain of the poorest members of society are not in the sample, but plenty of areas were sampled that were still relatively impoverished (e.g. Pilton, of Trainspotting fame, where I myself lived). Furthermore, I have already noted the existence of an age bias in the sample, with very many of the outright non-participants being over 60. However, it is difficult to see how a group of people who refused to even find out what the study was about could impart a bias to the data[1]. Selection biases usually apply when participants select themselves out for data relevant reasons. Perhaps it would be

[1] At a stretch one could argue that 'outright refusers' respond that way as a general defence against potentially traumatic situations, a defence which might have grown out of genuine trauma in earlier or later years. However, I don't count this likely - senior citizens just seem to be scared of young men at the door.

safer to say that a test of the models developed herein would be particularly useful amongst the over 60's. The survey data could be biased in other ways. For example, surveying tended to be done between 7 and 10pm on weekdays, and would thus exclude those who are not available at these times (though these times are least likely to suffer this kind of problem). Alternatively, 7-10pm might be a time when certain classes of people are more likely to come to the door. I think in particular this might be true of housewives. There is a slight female bias in the sample (71) to male (45), and it could well be that housewives answer the door because their husband is eating dinner, putting his feet up, or watching the footie. In addition, there were some occasions when a man answered the door and accepted the questionnaire, but the wife completed on the grounds that she was more interested in the survey than the man. Fortunately this did not occur frequently, but it did occur only in the male to female direction. However, as far as representativeness is concerned, whatever small or large discrepancies exist, they have not prevented us from testing and confirming essentially the same basic model of paranormal belief development across three highly diverse samples. Thus, to cut a long story short, I don't think representativeness need worry us unduly, though I would wish to see these models (especially the later more complex ones) tested on larger samples from a general population.

3) Fantasy proneness is implicated in accounts of 'false-memory' syndrome, so perhaps the correlation of trauma with fantasy is a result of false memory (Briere and Conte, 1993; Zelig and Beidleman, 1981). When false memory caused cases are removed, the correlations with trauma would deteriorate, and the model would look far less impressive. There are a number of strong responses to this criticism (which has actually been expressed to me before);

i) False memory accounts are associated primarily with particular methods for eliciting childhood memories (especially hypnosis, see Zelig and Beidleman, 1981)- though false memories may arise spontaneously, they are most likely to occur when psychotherapists use memory retrieval methods like hypnosis. The questionnaire method of assessing childhood trauma is not implicated amongst these techniques because it does not readily prime or suggest to respondents that they should recover memories. Its response format is more akin to recognition memory than recall memory. However, the apparent paucity of false memories occurring using non-hypnotic observation techniques may be illusory. It may be that we are more attuned to false memories occurring in the use of hypnosis because as professionals we have been concerned with the possibility of therapist induced memories. False memories may simply go undetected using other types of observation. I submit that it certainly could be that my questionnaire responses have been distorted by false memories. However, when I think of the variables in this study that have played the more prominent role in determining childhood fantasy and paranormal experience, like parental divorce, isolation, loss of friends and family I just can't believe ordinary people could be mistaken so greatly about their conscious recall of these events. Nevertheless, I accept that by themselves the results supporting a link between trauma and fantasy in this thesis do not guarantee that cause runs from trauma to fantasy - the path could go the other way. Inevitably, results from studies using other methodologies are needed to clarify the direction of influence.

ii) False memories are primarily about sexual or extreme physical trauma. Practically no sexual abuse trauma has been reported by respondents in this thesis. Therefore, the correlates established between trauma and fantasy, and trauma and experience do not

appear to depend crucially upon those types of trauma most susceptible to a false memory critique. Indeed, the lack of sexual abuse trauma reported would tend to favour an interpretation of response censorship, rather than confabulation. Interestingly, as a direct result of work by Irwin (1992) and myself (Lawrence et al., 1995) French and Kerman (1996) have compared levels of paranormal belief for a group of people with known childhood trauma histories, with a matched group of people with no known childhood trauma history. They found a significant difference in reported level of paranormal belief, with higher scores in the trauma group. As we have seen, false memory or fantastic confabulation does not seem to be critical to our results given the nature of the trauma reported, but in any case French and Kerman's (1996) study shows that a relationship exists between *real trauma* and paranormal belief. Nevertheless, false memories may not just pertain to sexual and physical abuse trauma but also to false memories of a quite mundane nature, involving say false memories of having been shopping last week in Safeways. Certainly I have been aware of believing I have done something just because at some time in the past I thought strongly about doing it - but these occasions are rare in the extreme, and so lest I become a wholesale memory nihilist I will suppose that generally people do faithfully report their memories accurately.

4) *Response censorship might have prevented people with low paranormal belief from expressing high levels of trauma, due to the socially undesirable nature of sexual and physical abuse trauma.* I take this criticism seriously because from point 3 above we noted a paucity of sexual abuse trauma in our data, and I suggested that this could be due to response censorship (people might be worried that acknowledging the existence of a sexual trauma would lead to further deeper inquiries from the researcher, and so not bother to acknowledge trauma of which they are otherwise quite aware). I think

the key thing here is that it is social desirability which leads to response censorship. Response censorship, if it occurs, would lead to potentially different correlations between experience, trauma and fantasy. Fortunately, I included a 13 item measure of social desirability (a reduced version of the Marlowe-Crowne social desirability scale) in the survey questionnaire for the SPR and Edinburgh surveys. Using 9 as a clinical cut-off point for the existence of a social desirability response set, we can compare the correlations for the full sample scores with correlations for the reduced (non-social desirability response set) scores;

Table 3. Correlations	Edinburgh Survey Full Sample (N=129)			Edinburgh Survey Reduced Sample (N=105)		
	PES	EPBS	CF	PES	EPBS	CF
EPBS	0.71			0.71		
CF	0.48	0.37		0.49	0.38	
STCE	0.24	0.26	0.27	0.25	0.25	0.28

Table 4. Correlations	SPR Survey Full Sample (N=58)			SPR Survey Reduced Sample (N=47)		
	PES	EPBS	CF	PES	EPBS	CF
EPBS	0.51			0.55		
CF	0.25	0.21		0.27	0.27	
STCE	0.46	0.18	0.19	0.44	0.17	0.13

In Tables 3 and 4 we can directly compare the correlations between the variables for paranormal experience (PES), paranormal belief (Edinburgh scale), childhood fantasy and overall trauma (STCE). It is manifestly apparent that the removal of cases likely to suffer from high socially desirability does not affect the overall correlation found at all. Thus, modelling attempts based upon the reduced correlation matrices would show little change from their full sample counterparts. Response censorship due to social desirability does not seem to have affected the data. Lastly, the questionnaire included an explicit honesty statement

designed to reduce the level of dishonest response (Eysenck, Eysenck and Shaw, 1974).

5) *The large survey questionnaire has a set order of questions, with no attempt to counterbalance question orders. Fatigue and boredom effects may affect the results, especially for items that come later in the questionnaire.* Alas, the survey questionnaire does contain this flaw, and there is potentially a problem of boredom or fatigue effects setting in as the respondent tries to plough through the questions. However, the questionnaire was explicitly constructed in five sections, each of which was designed to take no more than 9 minutes to complete, and participants were instructed on the questionnaire cover to '...do one section each evening, at a time that suits you...'. If they followed these instructions (annoyingly no check was made for this) then fatigue and boredom effects would be minimised. Perhaps then we should be more cautious about models based on data which appears substantially later in the questionnaire (the purpose in life test, and Affectometer 2). This would apply most strongly to the personal consequences model, which I have identified as needing replication anyway.

6) *the problem of 'alternative models'* - a persistent problem with structural equation models is that equally acceptable models may exist (from a statistical perspective) for a given covariance structure which differ radically in the kind of interpretation to which they give rise. For example, it could be argued that latent trait models implying no particular causal structure more satisfactorily explain the data, especially given the potential problems with cross-sectional data and items with content overlap addressed below. The trouble with many alternative models is that they fit the data just as well but are theoretically meaningless. However, some models may exist which represent the data in this thesis just as well, if not better, than the reported models, but which imply radically different causal or non-causal structures. I can not claim to have tackled all possible models

7) *the problem of cross sectional data* - cross sectional designs like the ones presented in this thesis involve the collection of data from one moment in time. Though often convenient for the researcher, they do not represent the optimum way of assessing causal influences of the kind being investigated in this thesis. A better way of investigating causal processes is by conducting a longitudinal study in which the same participants are measured on at least three separate occasions over a number of years. In such circumstances the causal ordering of variables in a model is not in dispute, and researchers may also assess the reliability of their measures over this time period. In contrast in my own analyses I am assuming that people's present day memories of childhood trauma and fantasy are accurately caused by the actual childhood events. This assumption may be appropriate to a greater or lesser extent, and in all likelihood it is appropriate although we may reasonably expect some distortion to have crept in to recollections of long past childhood over the years. Whether the causal sequences remain valid when tested using other methodologies only time, and more research can tell. However, I present evidence later in the discussion which draws upon alternative sources of evidence in establishing a likely causal role for childhood trauma and fantasy in the establishment of paranormal belief and experience.

8) *the problem of self-report data* - questionnaires are a quintessential form of self-report data. The researcher relies totally upon the participants' good memory, degree of self-insight, and honesty in order to establish a valid set of data. If any one of these three things falters then the validity of the report is thrown in to question. On memory, we have already spent much discussion. As far as insight is concerned we can only hope that participants had insight enough in to their own mentality to answer the questions set them. Fortunately few of the questionnaires used in this thesis demand great insight. However, the items assessing purpose in life and well being do demand of people that they reflect accurately upon

their recent or long term outlook on life. In addition, where purpose in life is concerned some individuals simply may not spend much time wondering about their role or purpose, or life's meaning. To the extent that anyone lacks insight those scales may be effected. A potential solution to peoples lack of insight about themselves, is to have other people comment upon the target person's character. Thus, a researcher might survey not just my opinion about my purpose in life, but that of my wife's and my father's. I think both of them would provide support for the notion that I have a sense of purpose and meaning in my life, or at the very least that I know what I want to do with it. Data of this kind could easily be 'factored' in to future analyses of the purpose in life construct, and might help to overcome problems of low insight, or at least raise new problems of whose insight to trust! As far as participant honesty is concerned I think we have covered that in its most relevant sense already in the discussion of social desirability. However, one could raise the question of outright, pernicious lying. I think the greatest problem with this is that it is almost certain to be rare, and likely to be motivated by some strong reaction to the questionnaire. I don't believe that such a reaction is likely to have occurred with my own research.

9) *the problem of content overlap* - where scales contain items which overlap considerably in content high correlations are an unsurprising result. Moreover, the best statistical approach to scales that have such overlap may be to factor them together in any structural equation modelling attempt, for parsimony's sake if nothing else. Clearly the paranormal belief scale and paranormal experience scale have content overlap in this sense, so might we want to factor these together in to one variable? This may be especially indicated where the correlation between the two unfactored variables is particularly high, and when attenuation for unreliability is accounted for then the urge to factor highly correlating scales may be irresistible. Thus, paranormal experiences regularly correlate about 0.70 with the latent variable of paranormal belief in the

structural equation models I have presented. Could we not just suppose that paranormal experience items are another facet of a wider 'paranormal issues' factor, rather than a separate and independent variable in their own right? I actually think that in this instance we should resist the temptation to factor together the paranormal variables. Firstly, I think the Edinburgh Paranormal Belief Scale is not typical of most paranormal belief scales in that it asks about one's friends psychic experiences, reading about psychic powers in newspapers, and psychic ability within the family. Other scales ask more straight-laced questions about ones belief in ESP per se, or PK per se, or life after death per se. For me the EPBS relates consistently more strongly to paranormal experience (than the ASGS-B) because it taps in to behaviours characteristic of the psychic experient (namely, that they have a circle of friends who accept their ability because of the friends' own psychic experience, or that the experient tends to be interested in reading about other psychic experiences, or that their family may also have individuals who have psychic ability). The EPBS draws perhaps too strongly on these items in its relation to paranormal experience, whereas the ASGS-B weighs in with a more typical 0.58 correlation with paranormal experience on its own. As I have noted in a previous chapter, there are many kinds of paranormal believer and not all of them will be psychic. The EPBS probably is biased towards those believers who are psychic. Secondly, from a conceptual perspective it is quite easy to conceive of individuals who hold high levels of paranormal belief but who have had little or no paranormal experience. Such individuals may be swayed by scientific evidence, or popular journalism, or occult literature, or the X-files(!), or the experience of friends and family. A 'paranormal views' factor which conglomerates paranormal belief and experience together as though one must necessarily give rise to the other does not do justice to the diversity of paranormal believers. Paranormal experience, whatever its ontological basis, undoubtedly gives rise to paranormal

belief, but paranormal belief by no means gives rise to paranormal experience. To factor variables like these together just because their content overlaps would prejudge issues of causal priority which are intuitively appealing and theoretically relevant. Lastly, if we feel justified in factoring together because the paranormal experience relates highly to a paranormal belief factor we should remember that this is a situation in which correction of correlations upwards for attenuation due to unreliability has already occurred for one half of the correlational relationship, namely the belief factor. The factor draws out the reliable systematic variance between the two belief variables, thus removing the unreliability that attenuates measured variables under normal circumstances. Thus, we might expect the correlation of 0.7 between experience and a belief factor to be pretty much at its peak - correction for attenuation would achieve little. However, I submit that there might be other areas of content overlap within the survey questionnaire, in particular between childhood fantasy and paranormal experience where questions about fairies and elves might sit thematically alongside concerns about psychic powers. Questions about content overlap accounting for these correlations can be easily answered by correlating non-overlapped versions of the paranormal experience and fantasy scales. I took all of the direct paranormally relevant items out of the childhood fantasy scale (items 4, 5, 6, 8 and 14) and recalculated the correlation with paranormal experience - it was 0.48, which is no less than the original correlation. Thus, content overlap, though it might be a problem in principle, seems not to be a problem in practice.

I have now outlined what I think are the most serious objections to the models presented in this thesis. None of them are fatal, as I see it, though some of them might yet have influenced the data in subtle ways which only further replication will reveal. But then, further replication would be needed whether there were flaws or not.

I think the models presented do portray a valid graphical narrative, caveats excepting, but for the models to be truly causal we need to make some appeal either to strong theoretical grounds for imputing causality, or to other evidence. In the remainder of this section on the causal models of paranormal belief and experience I want to briefly look over the autobiographies and biographies of some psychics, to see whether the themes coming through from the modelling research are echoed in psychics' developmental histories.

The Early Lives of the Psychics

In this thesis I have primarily been concerned with survey-based data pertaining to the prevalence of childhood trauma and fantasy in the development of paranormal experience and belief. As good as such techniques are for the demonstration of the strength of relationship between the modelled variables, we need to look elsewhere to provide wider validity to the model being built up. In short, if the model of paranormal belief development I am building has any real-world validity then this should be apparent in the lives of people who would call themselves psychic.

Fortunately there is no shortage of biographical material relating to the lives of psychics. In this section then I look over a selection of biographies of psychics, picked for no other reason than that those biographies were readily available (either in my own library or at the Koestler Chair library). I make no claim that the survey is representative, beyond saying that no obvious bias went in to the choice of biography. At the very least this chapter will help to enrich our understanding of the problem at hand. For some psychics I had to hand a reasonable amount of material, whilst for others I had only a little amount of biographical material. Just because I write a paragraph about one psychic, and many paragraphs about another, one shouldn't suppose there is any special relevance. I begin with an examination of the childhood of Eileen Garrett. Of all modern day psychics she is the one with whom I am most impressed.

The Childhood of Eileen Garrett

Eileen Garrett was a born in the first decade of this century in County Meath, Ireland. Whilst still a baby she was adopted by her Aunt and Uncle, and spent most of her childhood living with them. She seems to have had psychic ability from a very early age, and in her autobiography squarely places the cause, or part cause, of her talents down to the conducive folk legends of fairy kingdoms that were still commonplace amongst Irish folk at the turn of the century. However Garrett's abilities may have started she went on to become one of the most, if not the most, respected mental mediums of the twentieth century. Perhaps the single most striking thing about her mediumship is that she always maintained a critical distance from the phenomena she produced. Though spiritualistic beliefs provided a ready-made medium for her abilities, she never actually pronounced a strong belief in spiritualism. She readily submitted herself to scientific scrutiny on numerous occasions, and partly with money from her business career as the owner of a publishing company she founded the Parapsychological Foundation to support scientific research in to paranormal phenomena. Thus, many of the sceptics' arguments in favour of charlatanry in mediums simply fall flat in the face of Mrs. Garrett's life. There could be no monetary motive for her actions, for she made her money quite readily through her business. Fame was certainly not an interest of hers, for she already had the acquaintance of the good and great in the publishing world. It seems then that Mrs. Garrett provides a most interesting case - a medium whose integrity, at the least, is beyond question. All being equal, the details drawn from her childhood may provide revealing and valuable insights.

In her own autobiography, *Many Voices*, Garrett (1968) outlines in some detail her childhood in the 'magical realms' of Meath County, Ireland. Here she notes, the cultural context was steeped in old Irish legends of the fairy kings. One forms the distinct impression that persistent reference must have been made to tales of fairy and ghost for a child in the Ireland of the 1900's. Garrett recalls in particular being taken to Tara Hill by her Uncle Brownell, who

convinced her that the lands around the hill harboured 'mysterious sun temples' from the 'lost tribes' of the past days. A noted cousin also spoke of an ancient priest he had seen walking on Michaelmas Eve, this same cousin introduced the young Eileen to dowsing with a forked willow branch. As she wrote herself, 'Who knows how much the awed mind of a child, filled with wonder, absorbed from men such as these...' (pp. 18).

The above may suggest that Garrett's childhood was idyllic, a tale of those halcyon days in the emerald isle. The full story is somewhat less convivial though. Born to parents whose own personal strife ('confused with life and religion' in Garrett's own words) led to their committing suicide just after her birth, Eileen lived out the greater part of her childhood with her maternal Aunt and Uncle. Though she expressed much love for her Uncle, whose death during her childhood hurt her very much, she expresses only a grudging respect for her aunt. It seems that for most of her youth she simply clashed with her aunt. She describes going to bed early as a means of escape from her aunt (who would often send her there anyway) but beyond this punishment she seems to have found in her bedroom her own private world of imaginary friends and places. It is clear from Garrett's childhood that many of our childhood causes are not only evident, but explicitly acknowledged. Let us turn now to that most enigmatic of physical mediums, Daniel Dunglas Home.

Daniel Dunglas Home

It is interesting to note of Home, one of most famous mediums of the nineteenth century, that he suffered terribly from consumption as a child, and was sent to live with an aunt in Edinburgh (from the United States) on the advice of a doctor. Thus we have a loss of affectional bonds and isolation due to illness, as well as the fact of the illness itself (Braude, 1986).

Doris Stokes

The late Doris Stoke's autobiographies (*Voices in My Ear*, *More Voices in My Ear*, *Innocent Voices in My Ear*, *A Host of Voices*, and *Whispering Voices)*, sold in the

hundreds of thousands during the early eighties. I bought all of them as a lad. Unfortunately Doris did not have a particularly traumatic childhood, though of course living in Grantham, *and during a war*, life must have been hard.

Ann Walker

Author of the exotic[2], book *Little One: Message from Planet Heaven*, Walker (1994) says little explicitly about her childhood, but what she does say is highly revealing. Walker does not appear to have been psychic from an early age, like many mediums. Instead, 'During her darkest hour, at the age of thirty-seven, when she had believed she could no longer carry the burden of emotional pain that was a legacy from her brutal and traumatic childhood, White Arrow had appeared to her. With his help and guidance she was slowly able to rebuild her life and now, in turn, works full-time to help others overcome pain and confusion of their own.' (Inside back cover). I shall not elaborate.

Jenny Cockell

Cockell could not be regarded as a psychic as such. Certainly she would not define her life in that way. But nevertheless her book, *Yesterday's Children*, clearly marks her out as a woman of considerable sensitivity, as well as a person of high intelligence and integrity. In that book she provides details of her past life memory of 'Mary', an Irish woman who she felt she had been in the early part of this century. Having experienced dreams and images of life as Mary since she was small, Cockell ultimately embarks on a search for Mary and the children that Cockell felt sure Mary had left behind after her death. This search inevitably brings her to Malahide, a small village north of Dublin, where a Mary Sutton lived with her five children in the 1920's. Having identified her Mary as Mrs. Sutton, Jenny then goes in search of Mary's children, who are now in their fifties and sixties. Her search ultimately proves successful, and she is reunited with her past life family. Cockell's case has attracted some interest from parapsychologists. In a recent oral presentation of Cockell's case Mary

---

[2] Even by paranormal standards.

Rose Barrington reported her own investigation of the Cockell case. Having interviewed Cockell in depth, and those closely connected with Cockell's story, Rose Barrington then went on to interview more distant relatives in the family, on the reasonable assumption that they would be unlikely to be party to any deception. In addition. If Cockell's story is a conscious adult fabrication, then distant relatives would report no particular knowledge of her past life memories of Mary that Cockell claims occurred in childhood. Having interviewed an elderly aunt and uncle, Barrington found that their memory of Jenny's youth included events consistent with the notion that Mary was being spoken of in Jenny's childhood. This in itself only establishes that Cockell's story is not an *adult* fabrication; she could have course plotted the whole story from an early age for some unfathomably obscure reason. Nevertheless, Cockell's books do not convey the typical air of the psychic or medium. There is something about them which suggests their authenticity implicitly.

As for Cockell's autobiographical details, it would appear that she was a bright, intelligent child, growing up in a relatively normal household save for a rather doctrinaire and stern father. Her relationship with her father does seem to have been strained at times, but generally I can see no evidence for any clear pattern of trauma in her childhood.

Miscellaneous mediums

I have more recently become acquainted with the autobiographies of the late Tom Pilgrim, a spiritualist healer, and the late Leslie Flint, a more interesting exemplar of the mediumistic art of direct voice[3]. Pilgrim did not appear to have a particularly traumatic childhood, but Flint's autobiography begins thus, 'In spite of a childhood which would give any modern child psychiatrist nightmares, or perhaps because of it, I have reached the age of fifty-nine without falling prey to neurosis, psychosis or even the screaming meemies.' (p. 9). Interestingly, Flint notes that his mother gave birth to him out of wedlock, a real social scandal in those days, and Flint was born in a Salvation Army

[3] The production of audible voices in thin air, or from a seance trumpet.

Home in Hackney. However, Flint's father turns up, marries his mother, and that's when the trouble starts,

> *'Their marriage was doomed from the beginning. They were pitifully young, desperately poor, and they had been coerced in to maturity by their elders long before they were ready for it. [S]omehow they survived three years of discord, fighting each other every inch of the way until the 1914-18 war broke out when I was three years old. My father was one of the first men in St. Albans to volunteer for the army...simply to get away from the domestic hell he lived in.' (pp. 10).*

After the father left for the war, Flint's mother got a job in a munitions factory, and with the money she would make from it would often spend her nights out on the town with local men, even leaving young Flint on his own in the house (aged only three or four) late at night, as Flint recalls, 'One of my earliest memories, vivid to this day, is of waking up in the dark to find I was alone in the room which was our home and of feeling scared and lonely and desperately frightened that my mother had gone away like my father had gone and neither of them would ever come back.' (pp. 11). Eventually, Flint's mother eloped with another man and 'disappeared from my life'. He went to live with his grandmother where she inculcated strong principles in to the young Flint with a 'leathery hand', but nevertheless Flint respected his grandmother and it was at the age of seven whilst still living with her that he had his first psychic experience (seeing two soldiers come in to their kitchen following his Aunt Nell, who had just lost her husband in France). And so he went on to become a medium. Short of illness and long-term isolation, Flint's childhood seems to show many of the kinds of traumas that have been appearing in our survey data as significantly related to paranormal experience.

So much for the lives of the psychics about whom I have read. I don't claim my reading is representative, but I wonder how many other 'professions' would provide a similar showing for trauma in their list of autobiographies. It seems to me that, thought not every psychic has a traumatic past, these autobiographies are interesting. Of additional interest, the one psychic that I personally know very well has a traumatic childhood past to match the very

worst of childhoods, and that person certainly does think that the trauma they experienced had a part to play in their development of psychic ability. Then again I might yet make the acquaintance of ten more psychics for whom childhood was all Ribena and Sticklebricks.

I now wish to move on and consider briefly the findings from the two other areas of the thesis, my critique of the PBS, and the meta-analysis of the sheep-goat effect. I won't dwell to long on either of these areas because I think that the issues are fairly straightforward.

The measurement of paranormal belief

I think my critique of the PBS strongly suggested the existence of problems with that scale, sufficient to warrant construction of a new scale. However, that critique, in itself, could be passed off as mere armchair criticism, perhaps accurate in principle but perhaps also inaccurate in reality. Tobacyk (1995b) suggested that someone conduct a confirmatory factor analysis to test his view of the PBS and mine. I did this, and I think it's fair to say that Tobacyk's position is now also empirically bankrupt. My own initial five factor solution isn't perfect, but it is a lot less imperfect than Tobacyk's, and the claim that the PBS was oblique and not orthogonal was clearly demonstrated. Short of saying that the British samples used were not representative of American samples (in which case perhaps we should call the PBS the American PBS), I can not see any way of escaping the conclusion that a new measure of paranormal belief is needed. Although other measures do exist (e.g. Jones et al., 1977) I think they are actually uniformly worse than the PBS. For me, at the present time only Thalbourne and Delin's (1993) Australian Sheep-Goat Scale has any merit, and even there I am growing increasingly doubtful about its future.

The solution, that I expressed at length in chapter four, is to develop a new scale. I won't labour the points raised in that chapter again, but I can see a useful Ph.D. of the future centred around the attempt to build *a paranormal*

beliefs and influences scale. I should add that I think in many ways the critique of the PBS has been as important as the main empirical component of this thesis, and I have rarely seen a paper by a researcher within the field now who does not stop to qualify their use of the PBS with reference to Lawrence (1995a).

The Meta-analysis of the Sheep-Goat Effect

I cut my teeth on this paper, though it comes at the end of my thesis, and so it has now been in the public domain for four years. It seems to have been a useful paper, as it has citations numbering in excess of twenty at the last count. My thoughts on the import of the data differ now from my originally published thoughts. I once thought, as I think many parapsychologists do still, that the sheep-goat effect tells us something genuinely causal about belief and its effect upon scoring in ESP tests. For the most part I now think the sheep-goat effect tells us merely that there is an *association* between belief and scoring in ESP tests. This is ironic, as I did the meta-analysis so I could 'fix' at least one genuinely causal path in to the model of paranormal belief being developed. I discussed my reasons for thinking this at length in chapter 10, and I also discussed the kinds of studies that I think do get us closer to seeing the effect of belief upon ESP scoring, the Lovitts' (1981) design. What I would like to see in future work with the sheep-goat effect is a) some new work! (the field seems to have died a death), b) work using the Lovitts' design (or Lawrence and Norman's (1997) modification of it), c) more work on the sheep-goat effect in free-response settings. If researchers switch to examining the sheep-goat effect using Lovitts' design then they will have to treat the sheep-goat effect as an interesting effect in its own right, which is proper, rather than tacking it on as an opportunistic analysis in some other study as has tended to happen. I very much hope they do, because I think the meta-analysis at least shows that there is something going on.

Suggestions for future research

I'd like to think that my Ph.D. research hasn't just made a few hours work for the print office. The critique of the PBS, and soon the factor analysis of the PBS (Lawrence, Roe, and Williams, 1997), have and will change many researchers' opinions of that scale. Until a new scale is devised I suspect Thalbourne and Delin's (1993) scale will form the stop-gap for many, but in the long run a new scale is essential, if for no other reason than researchers will need it to replicate my other work!

As for the modelling research reported here. I know for sure that two researchers are following up my work already as reported in Lawrence et al. (1995). Harvey Irwin, whose work formed the basis for my thesis, is one of those attempting to follow up my work, and he has purchased a copy of EQS version 5.0 to do it! When I heard that I felt that I must have done something worthwhile! Simon Sherwood, a post-graduate at the Koestler Chair of Parapsychology, gained an Economic and Social Research Council grant on the basis of a proposal to follow up my work, once again using covariance structure modelling. Clearly other people think that work of this general kind is useful also.

The models I have examined here at the very least deserve further analysis. They seem to me to have a great deal of potential for further research. One thing that I would like to see done is a study to look at critical stages for the development of paranormal experience as a result of trauma. Bryant (1995) has shown that people who were sexually abused as children below the age of seven report significantly more fantasy proneness than those traumatised between seven and 14, who in turn report more fantasy proneness than those never having reported sexual abuse. One can explain Bryant's findings in a number of ways (e.g. false memory resulting from fantasy proneness), but it seems to me that the results can be explained in terms of trauma causing fantasy proneness with differential effect across the span of childhood development. That the age of below seven coincides with certain broadly

accepted cognitive developmental stages is interesting (Piaget's preoperational period from 2-7, when representational, i.e. imaginally derived, thought takes root). In any case, it seems to me that one might go on to ask whether the link between trauma and fantasy and therefore paranormal experience is one which derives only from a particularly sensitive stage in childhood like the period from 2-7. I think future research may profitably explore this possibility. Furthermore, I have thought throughout my modelling research that the paranormal experience - trauma link is similar to the dissociative/psychopathology - trauma link (e.g. Putnam, Guroff, Silberman, Barban and Post, 1986). Of course we know that magical ideation is related to paranormal belief (e.g. Thalbourne, 1985). I wonder whether the dissociative coping mechanisms used to deal with trauma reside along a continuum of severity. Thus, mild trauma requires almost no coping, medium level trauma perhaps elicits fantasy-based modes of coping (and thus paranormal experience), and severe trauma elicits more psychopathological modes of coping (e.g. multiple personality disorder). Researchers might want to look for weak quadratic curves in the relation between paranormal experience, fantasy and trauma as a way of shedding light on this possibility. To explore the paranormal experience/trauma link in a different context one might conduct a series of interviews with psychics. Likewise, if one had access to a substantial amount of cash one could conduct a longitudinal study of trauma and paranormal experience throughout childhood. This study would bring with it its own difficulties, but it might just resolve the question of whether trauma causes experience or fantasy 'causes' trauma and experience, at least for those kinds of trauma susceptible to a false memory analysis.

For my own part I think that future research in to the link between trauma and experience will need to be conducted by people working within the clinical realm (counsellors, clinical psychologists), especially if sexual abuse and physical abuse trauma are probed in to more deeply. Deeper probing increases the risk

of abreactive responses, and these are things that only trained personnel should deal with.

However, the personal consequences model offers much ground for further work for anyone who may be interested. In particular the apparent sex difference would seem to offer a fruitful starting place for future research. Why is it that men respond so differently, or weakly, to paranormal experience? Interviews with men and women who have experienced causal psychic events for short episodes in their lives may be instructive here. A content or discourse analysis on such material may help us see whether men characterise such experiences as threatening, whilst women see them as integrative episodes that enrich experience. These are interesting possibilities that extend beyond the modelling framework developed in this thesis. And what of that framework?

I very much hope my use of covariance structure modelling has been viewed as an advantage rather than a disadvantage in this thesis. I deliberately set out to show the methodological utility of this technique in the analysis of correlational data. At the very least no one can accuse me of having analysed my data in a piecemeal or unorganised manner, and that advantage alone suggests to me that covariance structure modelling has been worthwhile. Moreover, the analyses have always been couched within an explicit theoretical framework, and no one could now say they don't know what my perspective on paranormal belief is. I think the clarity of theoretical exposition that comes from covariance structure analysis will enable others to replicate my work, and extend it in to new areas, with relative ease.

At the very least that is *my* belief.

Adorno, T.W., Frenkel-Brunswick, E., Levinson, D.J. and Sandford, R.N. (1950). *The authoritarian personality.* New York: Harper and Row.

Aikaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52, 317-332.

Akers, C. (1987) Parapsychology is science, but it's findings are inconclusive. *Behavioural and Brain Sciences*, 10, 566-568.

Alcock, J.E. (1981). *Parapsychology: Science or Magic? A Psychological Perspective.* Elmsford, NY: Pergamon Press.

Alcock, J.E. and Otis, L.P. (1980). Critical thinking and belief in the paranormal. *Psychological Reports*, 46, 479-482.

Anderson, J. (1988). *Trance mediums and schizophrenics: A psychological assessment and P300 analysis.* Unpublished master's thesis, University of New England, Armidale, Australia.

Andrews, F.M. and Withey, S.B. (1976). *Social indicators of well being.* New York: Plenum.

Bainbridge, W.S. (1978). Chariots of the gullible. *Skeptical Inquirer, 3(2)*, 33-48.

Barnett, V. and Lewis, T. (1978) *Outliers in statistical data.* New York: John Wiley and Sons.

Beloff, J. (1963). Explaining the paranormal. *Proceedings of the Society for Psychical Research*, 42, 101-114.

Bentler, P. (1989) *EQS structural relations program manual.* Los Angeles, CA: BMDP Statistical Software.

Bhadra, B.H. (1966) The relation of test scores to belief in ESP. *Journal of Parapsychology*, 30, 1-17.

Biddle, B.J. and Marlin, M.M. (1987). Causality, confirmation, credulity, and structural equation modeling. *Child Development*, 58, 4-17.

Blackmore, S.J. (1988). Do we need a new psychical research? *Journal of the Society for Psychical Research*, 55, 49-59.

Blackmore, S.J. and Troscianko, T. (1985). Belief in the paranormal: Probability judgements, illusory control, and the "chance baseline shift". *British Journal of Psychology*, 76, 459-468.

Bollen, K.A. (1989). *Structural equations with latent variables.* John Wiley: New York.

Bookstein, F.L. (1984). The elements of latent variable models: A cautionary lecture. *Unpublished manuscript,* Center for Human Growth and Development, University of Michigan.

Braude, S.E. (1978). *ESP and Psychokinesis: A philosophical examination.* Philadelphia: Temple University Press.

Braude, S.E. (1986). *The limits of influence.* Routledge and Kegan Paul: London.

Briere, J. and Conte, J. (1993). Self-reported amnesia for abuse in adults molested as children. *Journal of Traumatic Stress*, 6, 21-31.

Broad, C.D. (1953) The relevance of psychical research to philosophy. In J. Ludwig (Ed.) *Philosophy and Parapsychology.* Buffalo: Prometheus.

Broughton, R. (1991) Repetition avoidance in ESP tests: Do sheep and goats really differ? *Proceedings of Presented Papers: The 34th Annual Convention of the Parapsychological Association*, 429-434.

Bryant, R.A. (1995). Fantasy proneness, reported childhood abuse, and the relevance of reported abuse onset. *International Journal of Clinical and Experimental Hypnosis*, 43, 184-193.

Byrne, B. M. (1994). *Structural equation modelling with EQS and EQS/Windows: Basic concepts, applications, and programming.* Sage: London.

Cattell, R.B. (1973) *Personality and mood by questionnaire.* San Fransisco: Jossey-Bass.

Cattell, R.B. and Jaspers, J. (1967) A general plasmode for factor analytic exercises and research. *Multivariate Behavioural Research Monographs, 3,* 1-212.

Cattell, R.B. and Vogelman, S. (1977) A comprehensive trial of the scree and KG criteria for determining the number of factors. *Multivariate Behavioural Research, 12,* 289-325.

Chamberlain, K. and Zika, S. (1988). Measuring meaning in life: An examination of three scales. *Personality and Individual Differences*, 9, 589-596.

Clarke, D. (1991) Belief in the paranormal: A New Zealand survey. *Journal of the Society for Psychical Research*, 57, 412-425.

Cliff, N. (1970) The relation between sample size and population characteristic vectors. *Psychometrika, 35*, 163-178.

Cockell, J. (1993). *Yesterday's Children.* Piatkus: London.

Cohen, D. and MacKeith, S.A. (1991). *The development of imagination: The private worlds of childhood.* Routledge: London.

Council, J.R. and Edwards, P.W. (1987) *Survey of traumatic childhood events.* Unpublished measure. Fargo: North Dakota State University.

Council, J.R. and Huff, K.D. (1990). Hypnosis, fantasy activity, and reports of paranormal experiences in high, medium and low fantasizers. *British Journal of Experimental and Clinical Hypnosis*, 7, 9-15.

Cronbach, L.J. and Meehl, P.E. (1955) Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.

Crumbaugh, J.C. (1968). Cross-validation of Purpose in Life test based on Frankl's concepts. *Journal of Individual Psychology*, 24, 74-81.

Crumbaugh, J.C. and Maholick, L.T. (1964). An experimental study in existentialism: the psychometric approach to Frankl's concept of noogenic neurosis. *Journal of Clinical Psychology*, 20, 200-207.

Davies, M.F. (1988). Paranormal beliefs in British and southern USA college students. *Psychological Reports*, 62, 163-166.

Davies, M.F. and Kirby, H.E. (1985). Multidimensionality of the relationship between perceived control and belief in the paranormal: Spheres of control and types of paranormal phenomena. *Personality and Individual Differences*, 6, 661-663.

Davis, G.A., Peterson, J.M. and Farley, F.H. (1974). Attitudes, motivation, sensation seeking, and belief in ESP as predictors of real creative behaviour. *Journal of Creative Behaviour*, 8, 31-39.

Dawes, R.M., Landeman, J. and Williams, J. (1984) Reply to Kurosawa. *American Psychologist.* 39, 74-75.

Eckblad, M. and Chapman, L.J. (1983). Magical ideation as an indicator of schizotypy. *Journal of Consulting and Clinical Psychology*, 51, 215-225.

Edge, H.L., Morris, R.L., Palmer, J. and Rush, J.H. (1986) *Foundations of Parapsychology*, London: Routledge and Kegan Paul.

Ellason, J.W. and Ross, C.A. (1997). Childhood trauma and psychiatric symptoms. *Psychological Reports*, 80, 447-450.

Emme, E.E. (1940). Modification and origin of certain beliefs in superstition among 96 college students. *Journal of Psychology*, 10, 279-291.

Emmons, C.F. and Sobal, J. (1981). Paranormal beliefs: Testing the marginality hypothesis. *Sociological Focus*, 14, 49-56.

Eysenck, H.J. (1967). Personality and extrasensory perception. *Journal of the Society for Psychical Research*, 44, 55-71.

Eysenck, S.B.G., Eysenck, H.J. and Shaw, L. (1974). The modification of personality and lie scale scores by special 'Honesty' instructions. *British Journal of Clinical Psychological*, 13, 41-50.

Fancher, R.E. (1979). *The pioneers of psychology*. Norton: London.

Fisher, R.A. (1921). On the probable error of a coefficient of correlation deducted from a small sample. *Metron*, 1, 3-32.

Flint, L. (1971). *Voices in the dark*. Macmillan: London.

Frank, J.D. (1977). Nature and functions of belief systems: Humanism and transcendental religion. *American Psychologist*, 32, 555-559.

French, C. C. (1992). Factors underlying belief in the paranormal: Do sheep and goats think differently? *The Psychologist: Bulletin of the British Psychological Society*, 5, 295-299.

French, C. C. and Kerman, M.K. (1996). Childhood trauma, fantasy proneness, and belief in the paranormal. *Paper presented at the 1996 London Conference of the British Psychological Society at the Institute of Education*, University of London, 17-18 December.

Freud, S. (1959). Creative writers and daydreaming. In J. Strachey (Ed. And Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 9, pp. 141-153). London: Hogarth. (Original work published in 1908).

Furnham, A. (1990). The fakeability of the 16PF, Myers-Briggs and FIRO-B personality measures. *Personality and Individual Differences*, 11, 711-716.

Furnham, A. and Steele, H. (1993). Measuring locus of control: A critique of general, children's health and work related locus of control measures. *British Journal of Psychology*, 84, 443-479.

Garrett, E. (1968). *Many voices: The autobiography of a medium.* George Allen and Unwin Ltd.: London.

Gauld, A. (1968). *The founders of psychical research.* Routledge and Kegan Paul: London.

Glicksohn, J. (1990). Belief in the paranormal and subjective paranormal experience. *Personality and Individual Differences*, 11, 675-683.

Gray, T. (1987). Educational experience and belief in paranormal phenomena. In F. B. Harrold & R.A. Eve (Eds.), *Cult Archaeology and Creationism: Understanding Pseudoscientific Beliefs about the Past (pp. 21-33).* Iowa City: University of Iowa Press.

Gray, T. and Mill, D. (1990). Critical abilities, graduate education (Biology vs. English), and belief in unsubstantiated phenomena. *Canadian Journal of Behavioural Science*, 22, 162-172.

Grimmer, M.R. and White, K.D. (1990) The structure of paranormal beliefs among Australian psychology students. *Journal of Psychology*, 124, 357-370.

Hair, J.F., Jr., Anderson, R.E., Tatham, R.L. and Black, W.C. (1995*). Multivariate data analysis with readings.* Prentice-Hall, Inc. (International Editions): New Jersey.

Haraldsson, E. (1981). Some determinants of belief in psychical phenomena. *Journal of the American Society for Psychical Research*, 75, 297-309.

Haraldsson, E. (1985a). Representative national surveys of psychic phenomena: Iceland, Great Britain, Sweden, USA, and Gallup's multinational survey. *Journal of the Society for Psychical Research*, 53, 145-158.

Haraldsson, E. (1985b). Interrogative suggestibility and its relationship with personality, perceptual defensiveness and extraordinary beliefs. *Personality and Individual Differences*, 6, 765-767.

Haney, W. (1954) *Measurement of the ability to discriminate between inferential and descriptive statements.* Doctoral dissertation, Northwestern University.

Happs, J.C. (1987). Conceptual conflict over pseudoscience: A case study involving teacher trainees and their belief in water divining. *The Skeptic* (Australia), 7(2), 21-28.

Hays, C. (1974) *A methodological investigation of the threat index and the introduction of a short form.* Unpublished senior thesis, University of Florida.

Hilgard, J.R. (1970) *Personality and hypnosis: A study of imaginative involvement.* Chicago: University of Chicago Press.

Hilgard, J.R. (1974) Imaginative involvement: Some characteristics of the highly hypnotizable and nonhypnotizable. *International Journal of Clinical and Experimental Hypnosis, 22,* 138-156.

Honorton, C. and Ferrari, D.C. (1989) "Future Telling": A meta-analysis of forced-choice precognition experiments, 1935-1987. *Journal of Parapsychology,* 53, 99-139.

Honorton, C., Ferrari, D.C. and Bem, D.J. (1991) Extraversion and ESP performance: A meta-analysis and a new confirmation. *Research in Parapsychology* 1990. Metuchen, N.J.: The Scarecrow Press.

Horn, J.L. (1965) A rationale and test for the number of factors in factor analysis. *Psychometrika,* 30, 179-185.

Hu, L., Bentler, P.M., and Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin,* 112, 351-362.

Hufford, D.J. (1982). *The terror that comes in the night, an experience-centered study of supernatural assault traditions.* Philadelphia: University of Philadelphia Press.

Irwin, H.J. (1985a) A study of the measurement and correlates of paranormal belief. *Journal of the American Society for Psychical Research,* 79, 301-326.

Irwin, H.J. (1985b). Parapsychological phenomena and the absorption domain. *Journal of the American Society for Psychical Research,* 79, 1-11.

Irwin, H.J. (1989). Extrasensory experiences and the need for absorption. *Parapsychology Review,* 20(6), 9-10.

Irwin, H.J. (1990) Fantasy proneness and paranormal beliefs. *Psychological Reports, 66,* 655-658.

Irwin, H.J. (1991a) The reasoning skills of paranormal believers. *Journal of Parapsychology,* 55, 281-300.

Irwin, H.J. (1991b) A study of paranormal belief, psychological adjustment, and fantasy proneness. *Journal of the American Society for Psychical Research, 85,* 317-331.

Irwin, H.J. (1992) Origins and functions of paranormal belief: The role of childhood trauma and interpersonal control. *Journal of the American Society for Psychical Research,* 86, 199-208.

Irwin, H.J. (1993) Belief in the paranormal: A review of the empirical literature. *Journal of the American Society for Psychical Research*, 87, 1-39.

Irwin, H.J. (1994). Childhood trauma and the origins of paranormal belief: A constructive replication. *Psychological Reports*, 74, 107-111.

Jahoda, G. (1969). The psychology of superstition. Penguin: Harmondsworth.

Joesting, J. and Joesting, R. (1969). Torrance's Creative Motivation Inventory and its relationship to several personality variables. *Psychological Reports*, 24, 30.

Jones, W.H., Russell, D.W. and Nickell, T.W. (1977). Belief in the paranormal scale: An objective instrument to measure belief in magical phenomena and causes. *Journal Supplement Abstract Service, Catalog of Selected Documents in Psychology*, 7, 100 (MS. 1577). Washington D.C.: American Psychological Association.

Joreskog, K.G. and Sorbom, D. (1981). *LISREL V. User's Guide.* Chicago: National Educational Resources.

Kamman, R. and Flett, R. (1983). Affectometer 2: A scale to measure current level of general happiness. *Australian Journal of Psychology*, 35, 257-265.

Killen, P., Wildman, R.W. and Wildman, R.W. II. (1974). Superstitiousness and intelligence. *Psychological Reports*, 34, 1158.

Kline, P. (1986). *A handbook of test construction.* London: Methuen.

Kline, P. (1990). *Intelligence: The psychometric view.* London: Routledge.

Kline, P. (1993). *The handbook of psychological testing.* London: Routledge.

Krippner, S. (1962). Creativity and psychic phenomena. *Indian Journal of Parapsychology*, 4, 1-20.

Lawrence, T.R. (1992) Subjective random generations and the reversed sheep-goat effect: A failure to replicate. *European Journal of Parapsychology*, 8, 131-144.

Lawrence, T.R. (1993) Gathering in the sheep and goats. A Meta-analysis of forced choice sheep-goat ESP studies, 1947-1993. *Proceedings of Presented Papers: The 36th Annual Convention of the Parapsychological Association*, 75-86.

Lawrence, T.R. (1994) Parapsychology, Knowledge and Wisdom: The Philosophy of Science and the Science of Psi. *The SPR International Conference in Bournemouth, 1994.*

Lawrence, T.R. (1995a). How many factors of paranormal belief are there?: A critique of the PBS. *Journal of Parapsychology*, 58, 3-25.

Lawrence, T.R. (1995b). Moving on from the Paranormal Belief Scale: A Final Reply to Tobacyk. *Journal of Parapsychology*, 59, 131-140.

Lawrence, T.R., Edwards, C., Barraclough, N., Church, S. & Heatherington, F. (1995) Modelling childhood causes of paranormal belief and experience: Childhood trauma and childhood fantasy. *Personality and Individual Differences*, 19, 209-215.

Lawrence, T.R., Roe, C.A. and Williams, C. (1997). Confirming the factor structure of the paranormal beliefs scale: Big orthogonal seven or oblique five? *Journal of Parapsychology*, 61, 13-31.

Lawrence, T.R. and Norman, G. (1997). Reversing sheep and goat ESP scores using an auditory subliminal decoy: A replication of Lovitts (1981). *Unpublished manuscript.*

Lester, D., Thinschmidt, J.S. and Trautman, L.A. (1987). Paranormal belief and Jungian dimensions of personality. *Psychological Reports*, 61, 182.

Long, J.S. (1994a). Confirmatory factor models. In M. S. Lewis-Beck (Ed.), *Factor Analysis and Related Techniques*. Sage: London.

Long, J.S. (1994b). Covariance structure models: An introduction to LISREL. In M. S. Lewis-Beck (Ed.), *Factor Analysis and Related Techniques*. Sage: London.

Lundeen, G.E. and Caldwell, O.W. (1930). A study of unfounded beliefs among high-school seniors. *Journal of Educational Research*, 22, 257-273.

Lynn, S.J. and Rhue, J.W. (1988). Fantasy proneness: Hypnosis, developmental antecedents, and psychopathology. *American Psychologist, 43*, 35-44.

MacCallum, R.C., Roznowski, M. and Necowitz, L.B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, 111, 490-504.

MacCallum, R.C., Wegener, D.T., Uchino, B.N. and Fabrigar, L.R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin*, 114, 185-199.

Maller, J.B. and Lundeen, G.E. (1934). Superstition and emotional maladjustment. *Journal of Educational Research*, 27, 592-617.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57, 519-530.

Mardia, K.V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. Sankhya, B36, 115-128.

Marks, D. and Kammann, R. (1980). *The psychology of the psychic.* Buffalo, NY: Prometheus Books.

Meier, A. and Edwards, H. (1974). Purpose in life test: Age and sex differences. *Journal of Clinical Psychology*, 30, 384-386.

Messer, W.S. and Griggs, R.A. (1989). Student belief and involvement in the paranormal and performance in introductory psychology. *Teaching of Psychology*, 16, 187-191.

Moon, M.L. (1975). Artists contrasted with non-artists concerning belief in ESP: A poll. *Journal of the American Society for Psychical Research*, 69, 161-166.

Murphy, G. (1963). Creativity and its relation to extrasensory perception. *Journal of the American Society for Psychical Research*, 57, 203-214.

Murphy, K. and Lester, D. (1976). A search for the correlates of belief in ESP. *Psychological Reports*, 38, 82.

Myers, S.A. (1983). The Wilson-Barber inventory of childhood memories and imaginings: Children's form and norms for 1337 children and adolescents. *Journal of Mental Imagery, 7,* 83-94.

Nadon, R., Laurence, J. and Perry, C. (1987). Multiple predictors of hypnotic susceptibility. *Journal of Personality and Social Psychology*, 53, 948-960.

Neppe, V.M. (1983) Temporal lobe symptomatology in subjective paranormal experients. *Journal of the American Society for Psychical Research, 77,* 1-29.

Newmark, C., Frerking, P., Cook, L., and Newmark, L. (1973) Endorsement of Ellis's irrational beliefs as a function of psychopathology. *Journal of Clinical Psychology*, 29, 300-302.

Otis, L. and Alcock, J.E. (1982). Factors affecting extraordinary belief. *Journal of Social Psychology*, 118, 77-85.

Otis, L. and Kuo, E.C.Y. (1984). Extraordinary beliefs amongst students in Singapore and Canada. *Journal of Psychology*, 116, 215-226.

Padgett, V.R., Benassi, V.A. and Singer, B.F. (1981). Belief in ESP among psychologists. In K. Frazier (Ed.), Paranormal Borderlands of Science (pp. 66-67).Buffalo, NY: Prometheus Books.

Palmer, J. (1971) Scoring in ESP tests as a function of belief in ESP. Part I. The sheep-goat effect. *Journal of the American Society for Psychical Research*, 65, 373-408.

Palmer, J. (1972) Scoring in ESP tests as a function of belief in ESP. Part II. Beyond the sheep-goat effect. *Journal of the American Society for Psychical Research*, 66, 1-25.

Palmer, J. (1973b) ESP scoring as predicted from four definitions of the sheep-goat variable. *Research in Parapsychology* 1972, 37-39.

Pilgrim, T. (1982). *Tom Pilgrim: Autobiography of a spiritualist healer.* Sphere: London.

Planer, F.E. (1988) *Superstition.* New York, Buffalo: Prometheus Books.

Pasachoff, J.M., Cohen, R.J. and Pasachoff, N.W. (1970). Belief in the supernatural among Harvard and West African university students. *Nature*, 227, 971-972.

Paulhus, D. L. (1983). Sphere-specific measures of perceived control. *Journal of Personality and Social Psychology*, 44, 1253-1265.

Paulhus, D.L. and Van Selst, M. (1990). The Spheres of Control Scale: 10YR of research. *Personality and Individual Differences*, 11, 1029-1036.

Planer, F. E. (1988). *Superstition.* Buffalo, NY: Prometheus Books.

Polzella, D. J., Popp, R. J. and Hinsman, M.C. (1975). *ESP?* Paper presented at a meeting of the American Psychological Association, Chicago.

Putnam, E.W., Guroff, J.J., Silberman, E.K., Barban, L. and Post, R.M. (1986). The clinical phenomenology of multiple personality disorder: Review of 100 recent cases. *Journal of Clinical Psychology*, 47, 285-293.

Radin, D. I., McAlpine, S. and Cunningham, S. (1994). Geomagnetism and psi in the ganzfeld. *Journal of the Society for Psychical Research*, 59, 352-363.

Rao, P.V. Krishna (1992) Fantasy proneness, reports of paranormal experiences and ESP test performance. *Journal of Indian Psychology, 10*, 27-34.

Randall, T.M. (1990). Belief in the paranormal declines: 1977-1987. *Psychological Reports*, 66, 1347-1351.

Randall, T.M. and Desrosiers, M. (1980) Measurement of supernatural belief: Sex differences and locus of control. *Journal of Personality Assessment*, 44, 493-498.

Rees, W.D. (1971). The hallucinations of widowhood. *British Medical Journal*, 2 October, 4:37-41.

Reker, G.T. (1977). The purpose in life test in an inmate population: an empirical investigation. *Journal of Clinical Psychology*, 33, 688-693.

Rhue, J.W. and Lynn, S.J. (1987). Fantasy proneness and psychopathology. *Journal of Personality and Social Psychology*, 53, 327-336.

Rhue, J.W., Lynn, S.J., Henry, S., Buhk, K. and Boyd, P. (1990-1991). Child abuse, imagination, and hypnotizability. *Imagination, Cognition and Personality*, 10, 53-63.

Rokeach, M. (1960) *The open and the closed mind.* New York: Basic Books.

Roney-Dougal, S. (1984). "Occult" conference questionnaire. *Journal of the Society for Psychical Research*, 52, 379-382.

Rosenthal, R. (1984). *Meta-analytic Procedures for Social Research.* Beverly Hills: Sage.

Rosenthal, R. and Rubin, D.B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology*, 74, 166-169.

Rotter, J. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80, No. 609.

Sargent, C.L. (1981). Extraversion and performance in 'extra-sensory' perception tasks. *Personality and Individual Differences*, 2, 137-143.

Satorra, A. and Bentler, P.M. (1988a). *Scaling corrections for statistics in covariance structure analysis.* Los Angeles: UCLA Statistics Series #2.

Satorra, A. and Bentler, P.M. (1988b). Scaling corrections for chi-square statistics in covariance structure analysis. *Proceedings of the American Statistical Association*, 308-313.

Scheidt, R. (1973) Belief in supernatural phenomena and locus of control. *Psychological Reports*, 32, 1159-1162.

Schlitz, M.J. and Honorton, C. (1992). Ganzfeld psi performance within an artistically gifted population. *Journal of the American Society for Psychical Research*, 86, 83-98.

Schmeidler, G.R. (1943) Predicting good and bad scores in a clairvoyance experiment: A preliminary report. *Journal of the American Society for Psychical Research*, 37, 103-110.

Schmeidler, G.R. and McConnell, R.A. (1958) *ESP and Personality Patterns*. New Haven: Yale University Press.

Schoenberg, R. (1982). *MILS: A computer program to estimate the parameters of multiple indicator linear structural equation models*. Bethesda, MA: National Institutes of Health.

Schultz, W.C. (1957;1978). *FIRO Awareness scales manual*. Palo Alto, CA: Consulting Psychologists Press.

Schumaker, J.F. (1987). Mental health, belief deficit compensation, and paranormal belief. *Journal of Psychology*, 121, 451-457.

Schumaker, J.F. (1990). *Wings of illusion: The origin, nature, and future of paranormal belief*. Buffalo, NY: Prometheus Books.

Sheils, D. and Berg, P. (1977). A research note on sociological variables related to belief in psychic phenomena. *Wisconsin Psychologist*, 14, 24-31.

Singer, B and Benassi, V.A. (1981). Occult beliefs. *American Scientist*, 69, 49-55.

Smith, M.D., Foster, C.L. and Stovin, G. (1995). Intelligence and paranormal belief: The importance of context. *The Parapsychological Association 38th Annual Convention: Proceedings of Presented Papers*. pp. 444-453.

Sobal, J. and Emmons, C.F. (1982) Patterns of belief in religious, psychic, and other paranormal phenomena. *Zetetic Scholar*, No. 9, 7-17.

Stanford, R.G. (1984) Recent Ganzfeld-ESP research: A survey and critical analysis. In S. Krippner (Ed.) *Advances in Parapsychological Research*, Volume 4. pp. 83-112. Jefferson, N.C.: McFarland.

Stanovich, K.E. (1989). Implicit philosophies of mind: The Dualism Scale and its relation to religiosity and belief in extrasensory perception. *Journal of Psychology*, 123, 5-23.

Steering Committee of the Physicians Health Study Research Group. (1988) Preliminary report: Findings from the aspirin component of the ongoing physicians' health study. *The New England Journal of Medicine*, 318, 262-264.

Ter Keurst, A.J. (1939). Comparative differences between superstitious and non-superstitious children. *Journal of Experimental Education*, 7, 261-267.

Thalbourne, M.A. (1981) Extraversion and the sheep-goat variable: A conceptual replication. *Journal of the American Society for Psychical Research*, 75, 105-120.

Thalbourne, M.A. (1982) *A Glossary of terms Used in Parapsychology*. London: Heinemann.

Thalbourne, M.A. (1984). Some correlates of belief in psychical phenomena: A partial replication of the Haraldsson findings. *Parapsychology Review*, 15(2), 13-15.

Thalbourne, M.A. (1985). Are believers in psi more prone to schizophrenia? In R.A. White and J. Solfvin (Eds.), *Research in Parapsychology 1984*, pp. 85-88. Metuchen, NJ: Scarecrow Press.

Thalbourne, M.A. (1994). Belief in the paranormal and its relationship to schizophrenia-relevant measures: A confirmatory study. *British Journal of Clinical Psychology*, 33, 78-80.

Thalbourne, M.A. (unpublished manuscript) Further studies of the measurement and correlates of belief in the paranormal. Department of Psychology, University of Adelaide, Australia.

Thalbourne, M.A. and Haraldsson, E. (1980) Personality characteristics of sheep and goats. *Personality and Individual Differences*, 1, 180-185.

Thalbourne, M.A., Beloff, J., Delanoy, D.D. and Jungkuntz, J. (1983). Some further tests of the extraverted sheep versus introverted goats hypothesis [Summary]. In W.G. Roll, J. Beloff and R.A. White (Eds.), *Research in Parapsychology 1982* (pp. 199-200). Metuchen, NJ: Scarecrow Press.

Thalbourne, M.A., Dunbar, K.A., and Delin, P.S. (1989). An investigation into correlates of belief in the paranormal. Department of Psychology, University of Adelaide, Australia. *Unpublished manuscript.*

Thalbourne, M.A. and Delin, P.S. (1993). A new instrument for measuring the sheep-goat variable: Its psychometric properties and factor structure. *Journal of the Society for Psychical Research*, 59, 172-186.

Tobacyk, J.J. (1984). Paranormal belief and college grade point average. *Psychological Reports*, 54, 217-218.

Tobacyk, J.J. (1988) *A revised paranormal belief scale.* Unpublished manuscript. Louisiana Tech University, Ruston, LA.

Tobacyk, J.J. (1991) Superstition and belief about the prediction of future events. *Psychological Reports, 68*, 511-512.

Tobacyk, J.J. (1995a) What is the correct dimensionality of paranormal beliefs?: A reply to Lawrence's critique of the Paranormal Belief Scale. *Journal of Parapsychology*, 58, 27-46.

Tobacyk, J.J. (1995b). Final thoughts on issues in the measurement of paranormal beliefs. *Journal of Parapsychology*, 59, 141-145.

Tobacyk, J.J. & Milford, G. (1983) Belief in paranormal phenomena: Assessment instrument development and implications for personality functioning. *Journal of Personality and Social Psychology*, 44, 1029-1037.

Tobacyk, J.J., Miller, M. and Jones, G. (1984). Paranormal beliefs of high school students. *Psychological Reports*, 55, 255-261.

Tobacyk, J.J. & Mitchell, T.P. (1987) The out-of-the-body experience and personality adjustment. *Journal of Nervous and Mental Disease*, 175(6), 367-370.

Tobacyk, J.J., Miller, M., Murphy, P. and Mitchell, T. (1988). Comparisons of paranormal beliefs of black and white university students from the southern United States. *Psychological Reports*, 63, 492-494.

Tobacyk, J.J., Nagot, E., and Miller, M. (1988). Paranormal beliefs and locus of control: A multidimensional examination. *Journal of Personality Assessment*, 52, 241-246.

Tobacyk, J.J., Pritchett, G. and Mitchell, T. (1988). Paranormal beliefs in late-adulthood. *Psychological Reports*, 62, 965-966.

Tobacyk, J.J., Nagot, E., and Mitchell, T. (1989) Prediction of future events scale: Assessment of beliefs about predicting the future. *Journal of Social Psychology*, 129, 819-823.

Tobacyk, J.J. and Wilkinson, L.V. (1990). Magical thinking and paranormal beliefs. *Journal of Social Behaviour and Personality*, 5(4), special issue, 255-264.

Tobacyk, J.J. and Pirtilla-Blackman, A. (1992). Paranormal beliefs and their implications in university students from Finland and the United States. *Journal of Cross-Cultural Psychology*, 23, 59-71.

Tomic, van der Sijde, P.C. & Snel, F.(unpublished research) Factor analysis of the PBS on a large Dutch sample. (Details from Dr. P.C. van der Sijde, Transfer, Research, and Development, University of Twente. Holland.)

Tyrell, G.N.M. (1943). *Apparitions*. Society for Psychical Research: London.

Utts, J. (1991) Replication and meta-analysis in parapsychology. *Statistical Science*, 6, 363-403.

Valentine, W.L. (1936). Common misconceptions of college students. *Journal of Applied Psychology*, 20, 633-658.

Van de Castle, R.L. and White, R.R. (1955) A report on a sentence completion form of sheep-goat attitude scale. *Journal of Parapsychology*, 19, 171-179.

Wagner, M.W. and Ratzeburg, F.H. (1987). Hypnotic susceptibility and paranormal belief. *Psychological Reports*, 60, 1069-1070.

Walker, A. (1994). *Little one: Message from planet Heaven.* Century: London.

White, R.A. (1990). An experience-centered approach to parapsychology. *Exceptional Human Experiences*, 8, 7-36.

Wierzbicki, M. (1985) Reasoning errors and belief in the paranormal. *Journal of Social Psychology*, 125, 489-494.

Williams, L.M. and Irwin, H.J. (1991) A study of paranormal belief, magical ideation as an index of schizotypy and cognitive style, *Personality and Individual Differences*, 12, 1339-1348.

Wilson, I. (1995). *In search of ghosts.* Headline: London.

Wilson, S.C. and Barber, T.X. (1983) The fantasy prone personality: Implications for understanding imagery, hypnosis, and parapsychological phenomena. In A.A. Sheikh (Ed.), *Imagery: Current theory, research, and application* (pp.340-387). New York: Wiley.

Windholz, G. and Diamant, L. (1974). Some personality traits of believers in extraordinary phenomena. *Bulletin of the Psychonomic Society*, 3, 125-126.

Wuthnow, R. (1976). Astrology and marginality. *Journal for the Scientific Study of Religion*, 15, 157-168.

Zapf, R. (1976). Relationship between belief in superstitions and other factors. *Journal of Educational Research*, 38, 561-579.

Za'rour, G.J. (1972). Superstitions among certain groups of Lebanese Arab students in Beirut. *Journal of Cross-Cultural Psychology*, 3, 273-282.

Zika, S. and Chamberlain, K. (1992). On the relation between meaning in life and psychological well being. *British Journal of Psychology*, 53, 155-162.

Zelig, M. and Beidleman, W.B. (1981). The investigative use of hypnosis: A word of caution. *International Journal of Clinical and Experimental Hypnosis*, 29, 401-412.

Zuckerman, M. (1971) Dimensions of sensation seeking. *Journal of Consulting and Clinical Psychology*, 36, 45-52.

Zwick, W.R. and Velicer, W.F. (1982) Factors influencing four rules for determining the number of components to retain. *Multivariate Behavioural Research*, 17, 253-269.

Zwick, W.R. and Velicer, W.F. (1986) Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99, 432-442.

Zusne, L. and Jones, W.H. (1982). *Anomalistic psychology: A study of extraordinary phenomena of behaviour and experience.* Hillsdale, NJ: Lawrence Earlbaum Associates.

Meta-analysed Studies

Adcock, C.J. and Quatermain, D. (1959) Some problems of group testing of ESP. *Journal of Parapsychology*, 23, 251-256.

Ballard, J.A. (1988) Affective assessment, belief in ESP, and relaxation in relation to scoring on a psi task. *Research in Parapsychology* 1987, 51-55.

Banham, K.M. (1968) The imaginary quiz as a technique. *Journal of Parapsychology*, 32, 63-64.

Bevan, J.M. (1947) The relation of attitude to success in ESP scoring. *Journal of Parapsychology*, 11, 269-309.

Carpenter, J.C. (1971) The differential and hidden target differences consisting of erotic and neutral stimuli. *Journal of the American Society for Psychical Research*, 65, 204-214.

Casper, G.W. (1951) A further study of the relation of attitude to success in ESP scoring. *Journal of Parapsychology*, 15, 139-145.

Crandall, J.E. (1993) Psi-missing and displacement under relatively unfavourable testing conditions, using salient and non-salient targets. *Journal of the American Society for Psychical Research*, 87, 85-98.

Eilbert, L. and Schmeidler, G.R. (1950) A study of certain psychological factors in relation to ESP performance. *Journal of Parapsychology*, 14, 53-74.

Eisenbud, J. (1965) Perception of subliminal visual stimuli in relation to ESP. International *Journal of Parapsychology*, 7, 161-181.

Gerstein, P. and Merker, P. (1964) Attitudes toward ESP and choice of target as related to ESP performance. *Journal of Parapsychology*, 28, 226-227.

Hagio, S. (1989) An experimental research into the reliability of ESP scores in relation to personality traits. *Journal of Parapsychology*, 53, 88-89. (Actual data taken from original paper provided by Dr. Hagio).

Haraldsson, E. (1975) Reported dream recall, precognitive dreams, and ESP. *Research in Parapsychology* 1974, 47-48.

Haraldsson, E. (1976) Reading habits, belief in ESP and precognition. *Research in Parapsychology* 1975, 28-29.

Haraldsson, E. (1980) Scoring in a precognition test as a function of the frequency of reading on psychical phenomena and belief in ESP. *Research Letter* , No. 10, 1-8. (Published at the University of Utrecht Parapsychology Laboratory.)

Harary, S.B. (1977) A study of psi, memory and expectancy. *Research in Parapsychology* 1976, 121-126.

Honorton, C. (1972) Reported frequency of dream recall and ESP. *Journal of the American Society for Psychical Research*, 66, pp. 369-374.

Johnson, M. (1969) Attitude and target differences in a group precognition test. *Journal of Parapsychology*, 33, pp. 324.

Kahn, S.D. (1952) Studies in extra-sensory perception: experiments utilizing an electronic scoring device. *Proceedings of the American Society for Psychical Research*, 25, (Series 2,3 and 4).

Kennedy, J.E. and Haight, J. (1978) Are psychological tasks nonintentional psi tasks? *Journal of Parapsychology*, 42, 33-50. (Experiment. 2)

Lawrence, T.R. (1992) Subjective random generations and the reversed sheep-goat effect: A failure to replicate. *European Journal of Parapsychology*, 8, 131-144. (Prove ESP condition).

Lomax, L.F. (1987) Psi scoring as a function of private-public self consciousness and conflicting internal-external standards. *Proceedings of the Parapsychological Association: The 30th Annual Convention of the Parapsychological Association*, 171-183.

Lovitts, B.E. (1981) The sheep-goat effect turned upside down. *Journal of Parapsychology*, 45, 293-309. (Prove ESP condition).

McGuire, K., Percy, E. and Carpenter, J.C. (1974) A multivariate approach to the prediction of ESP test performance. *Research in Parapsychology* 1973, 34-35.

Musso, J.R. (1965) ESP experiments with primary school children. *Journal of Parapsychology*, 29, 115-121.

Nash, C.B. (1965) Opposite scoring direction in ESP. *Journal of Parapsychology*, 29, 122-126.

Osis, K. and Dean, D. (1964) The effect of experimenter differences and subjects belief upon ESP score. *Journal of the American Society for Psychical Research*, 58, 158-185.

Palmer, J. (1977) A group precognition study with college students. *Journal of Parapsychology*, 41, pp. 237.

Palmer, J. (1973b) ESP scoring as predicted from four definitions of the sheep-goat variable. *Research in Parapsychology* 1972, 37-39.

Palmer, J. and Miller, A. (1970) Monetary incentive and the sheep-goat effect. Proceedings of the Parapsychological Association: The 13th Annual Convention of the Parapsychological Association. 11-12.

Petersen, C.J. (1971) *Relationships among attitude, memory and extrasensory perception.* Unpublished thesis (M.Sc. dissertation- Wisconsin State University- August 1971).

Ryzl, M. (1968a) Precognition scoring and attitude towards ESP. *Journal of Parapsychology*, 32, 1-8.

Ryzl, M. (1968b) Precognition scoring and attitude. *Journal of Parapsychology*, 32, 183-184.

Schmeidler, G.R. and McConnell, R.A. (1958) *ESP and Personality Patterns.* New Haven: Yale University Press.

Schmeidler, G.R. (1962) Tests of creative thinking and ESP scores. Indian *Journal of Parapsychology*, 4, 51-57.

Schmeidler, G.R. (1964) An experiment on precognitive clairvoyance. Part V. Precognition scores related to feelings of success. *Journal of Parapsychology*, 28, 109-125.

Smith, K. and Canon, H.J. (1954) A methodological refinement of the study of

ESP, and negative findings. *Science*, 120, 148-149. (Leftward targets).

Taddonio, J.L. (1975) Attitudes and expectancies in ESP scoring. *Journal of Parapsychology*, 39, 289-295.

Taddonio, J.L., O'Brien, J., MacDonald, S. and Bibeau, J. (1979) ESP in life and lab: An investigation of intuition and card calling. *Research in Parapsychology* 1978, 86-87.

Taetzch, R.L. (1964) Sheep-goat and decline results of computer scored audience responses. *Journal of Parapsychology*, 28, 277-278.

Thalbourne, M., Beloff, J. and Delanoy, D. (1982) A test for the "extraverted sheep versus the introverted goats" hypothesis. *Research in Parapsychology* 1981, 155-156.

Thalbourne, M., Beloff, J. and Jungkuntz, J. (1983) Some further tests of the extraverted sheep versus introverted goats hypothesis. *Research in Parapsychology* 1982, 199-200.

Thorisson, K.R., Skulason, F. and Haraldsson, E. (1991) Effects of belief in ESP and distorted feedback on a computerised clairvoyance task. *Journal of Parapsychology*, 55, 45-58.

Van de Castle, R.L. and White, R.R. (1955) A report on a sentence completion form of sheep-goat attitude scale. *Journal of Parapsychology*, 19, 171-179.

Van de Castle, R.L. (1957) Differential patterns of ESP scoring as a function of differential attitudes towards ESP. *Journal of the American Society for Psychical Research*, 51, 43-61.

Woodruff, J.L. and Dale, L.A. (1950) Subject and experimenter attitudes in relation to ESP scoring. *Journal of the American Society for Psychical Research*, 44, 87-112.

Appendix I

Relevant Survey Questionnaire Items

In this appendix I provide basic details for those scales which have been constructed specially for this thesis. Though many other scales were included in the large survey questionnaire (see Appendix II) I have not used them in the thesis research, and thus do not dwell upon them.

Paranormal Experiences Scale

A principle factor analysis of this scale shows that it is unidimensional, with the single factor accounting for around 40% of the variance in items (construct reliability[1] is 0.90). Corrected Item-total correlations are given in brackets, showing that almost all items correlate highly with the unbiased total score. Items are scored on a seven point scale from strongly disagree (1) to strongly agree (7).

1. I have psi ability. (0.63)
2. At least once, I have been psychically aware of another's thoughts. (0.68)
3. I am often able to find people's lost objects very easily. (0.51)
4. At least once I have successfully seen the future. (0.73)
5. At least once, I have been able to move objects with my mind alone. (0.35)
6. I have experienced visions for which there were no normal causes. (0.78)
7. At least once, I have experienced my 'self' being separate from my body. (0.55)
8. I have experienced telepathy with another person. (0.54)

[1] Coefficient Alpha, calculated according to Kline (1993) formula 3.8.

9. I have experienced a future event before it happened. (0.72)
10. I have 'seen' or 'heard' the spirits of deceased people, friends and relatives. (0.53)
11. At times objects that I knew could only be in one location, went missing and were unaccountably found in another location. (0.52)
12. I am sometimes able to influence people, and objects just by thinking about them. (0.61)
13. I am psychic. (0.62)
14. I possess marked psychokinetic ability. (0.49)

Edinburgh Paranormal Belief Scale

This 9 item scale was thought up by the author as a complement to the ASGS scale. Principal factor analysis shows that a unidimensional solution is appropriate, although perhaps lacking in explanatory power (variance extracted = 31%, construct reliability = 0.8). Item-total correlations are acceptable for all but the last item. Items are scored on a seven point scale from strongly disagree (1) to strongly agree (7).

1. ESP is a real human ability. (0.53)
2. Friends of mine often tell me about their own psychic experiences. (0.57)
3. I often read about psychic phenomena in the newspapers. (0.48)
4. I often read about other people's ability to 'see' in to the future. (0.54)
5. Psychics can often move objects with the power of the mind alone. (0.56)
6. I often hear about psychokinetic events from friends and relatives. (0.50)
7. Certain members of my family have strong psychic ability. (0.51)

8. I could demonstrate ESP under controlled conditions. (0.43)
9. I could demonstrate PK under controlled conditions. (0.30)

Childhood Fantasy Scale

This 17 item scale is an adaptation of Myers' (1983) ICMI:C, and takes only those questions which refer to childhood fantasy, or the ability to recollect childhood memories. It is completed using the true/false response format. A principle factor analysis with direct oblimin rotation of the scale as used here shows that it is composed of two factors accounting for about 30% of the variability (construct reliability[2] = 0.74). The first factor loads strongly on items referring to pretence or make-believe, whilst the second loads on childhood reminiscence and belief in fantasy creatures and realms. However, these two factors are best represented as oblique and the correlation between them is reasonably high at 0.38. Future research should be done to examine the differential correlative power of these two subscales.

1. I can remember clearly one or more things that happened to me when I was three years of age or younger. (0.43)
2. When I remember back to when I was six or seven years of age, I can 'see' and 'hear' again what I saw and heard then, and can feel again the feelings (happy, sad, afraid, etc.) I felt then. (0.60)
3. Although I have grown, I still feel almost the same as I did when I was younger. (0.39)
4. When I was younger, I believed in such things as elves, witches, leprechauns, fairies, etc. (0.55)
5. When I was younger, I would dream or make believe I was flying so clearly that *I felt as if I really did fly.* (0.38)
6. When I was younger, I enjoyed fairy tales. (0.35)

---

[2] Calculated using the Kuder-Richardson 20 formula (3.9) given in Kline (1993), p. 40.

7. When I was younger, I was *very good* at make believe and imagining. (0.52)
8. When I was younger, I spent time thinking about such things as the meaning of life and death. (0.50)
9. When I was a young child (below age 8) I liked playing make-believe games such as cowboys, school teacher, husband/wife, etc. I like them better than games without make-believe such as draughts, building things, ball games, hopscotch, etc. (0.42)
10. When I was playing make-believe games as a young child, I would make-believe so well that what I pretended seemed real to me. (0.45)
11. When I was younger, I lived in a make-believe world *much or most of the time.* (0.47)
12. When I was *much younger*, I believed that my doll(s) or stuffed animal(s) were alive and had feelings (that is, they could feel pain, lonely, happy, etc.). (0.54)
13. When I was younger, I had a pretend friend or companion (such as a make-believe person, animal, or object) which I talked to, or took along with me. (0.29)
14. When I was younger, I would at times pretend and in some ways believe I was someone else such as a prince, princess, Snow White, Peter Pan, etc. (0.43)
15. When I was younger, I would have liked to take ballet lessons or I do or did *enjoy* taking ballet lessons. (0.20)
16. When I was younger, I would spend at least half of the time I was awake pretending or imagining. (0.59)
17. At times, when I was younger, it was hard for me to tell if something had actually happened or if I had *imagined* it happened. (0.21)

Appendix II

The Edinburgh and SPR Survey Questionnaire

Belief in the Paranormal
Survey Questionnaire

Dear Participant

The results of this survey will make a significant contribution to the study of belief in the paranormal, and paranormal experiences. If you are very sceptical, and are inclined to view this questionnaire as a waste of time, please do not. We need to know as much about people's reasons for scepticism as we do for those who believe.

The questionnaire takes about 45 minutes to complete- **but I have split it in to five convenient sections for you to fill in, that will not take more than 9 minutes each.** This way you could do one section every evening, at a time that suits you, and have it all finished in a week. Enclosed with this questionnaire is a separate test booklet, the *FIRO:B*, which I would also like you to fill in and return with this questionnaire. Psychologists are very aware that when people fill in questionnaires such as this, they sometimes give answers which they feel present them in a better light. For example, one of the questions might be: "Have you ever told a lie?"- the obvious answer to this is "Yes", but we still find that there are a number of people who say "No". Therefore the questionnaire includes some items to check those people on this. Remember that there are **no right or wrong answers**- this is not a test of ability- so please fill in the questionnaire as honestly as possible.

Thank you for your assistance.

Section 1. (Takes approx. 9 minutes)

I.D.:
Date of Birth:
Sex: M / F
Age:
Marital Status: Widower(ee) / Single / Married / Divorced

Listed below are a number of statements. Each statement is accompanied by a seven point scale. You must indicate to what extent you agree, or disagree, with each statement, *in so far as it applies to you.* Please circle a number for each statement, according to the following scale;

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Slightly Disagree	Uncertain	Slightly Agree	Moderately Agree	Strongly Agree

1. Some of my most vivid memories are called up by scents and smells.

1.....2.....3.....4.....5.....6.....7

2. Sometimes thoughts and images come to me without the slightest effort.

1.....2.....3.....4.....5.....6.....7

3. My thoughts often don't occur as words but as visual images.

1.....2.....3.....4.....5.....6.....7

4. If I wish I can imagine that my body is so heavy that I could not move it if I wanted to.

1.....2.....3.....4.....5.....6.....7

5. I am able to wander off into my thoughts while doing a routine task and actually forget that I am doing the task, and find a few minutes later that I've completed it.

1.....2.....3.....4.....5.....6.....7

6. If I wish I can imagine (or daydream) some things so vividly that they hold my attention as a good movie or story does.

1.....2.....3.....4.....5.....6.....7

7. I can sometimes recollect certain past experiences in my life with such clarity and vividness that it is like living them again or almost so.

1.....2.....3.....4.....5.....6.....7

8. If I stare at a picture and then look away from it, I can sometimes see an image of the picture almost as if I were still looking at it.

1.....2.....3.....4.....5.....6.....7

Please bear in mind the following basic definitions when completing this section;

Psi- general psychic ability, split into;

ESP- psychic reception of information without the use of known senses or reasoned guesses.

PK (Psychokinesis)- psychic influence on physical objects by force of mental will power alone.

9. I have Psi ability.

1.....2.....3.....4.....5.....6.....7

10. ESP is a real human ability.

1.....2.....3.....4.....5.....6.....7

11. At least once, I have been psychically aware of another's thoughts.

1.....2.....3.....4.....5.....6.....7

12. Friends of mine often tell me about their own psychic experiences.

1.....2.....3.....4.....5.....6.....7

13. I am often able to find other peoples lost objects very easily.

1.....2.....3.....4.....5.....6.....7

14. I often read about psychic phenomena in the papers.

1.....2.....3.....4.....5.....6.....7

15. At least once, I have successfully 'seen' the future..

1.....2.....3.....4.....5.....6.....7

16. I often read about other people's ability to 'see' the future.

1.....2.....3.....4.....5.....6.....7

17. Psychics can often move objects with the power of the mind alone.

1.....2.....3.....4.....5.....6.....7

18. At least once, I have been able to move an object with my mind alone.

1.....2.....3.....4.....5.....6.....7

19. I often hear about psychokinetic events from friends and relatives.

1.....2.....3.....4.....5.....6.....7

20. Certain members of my family have strong psi ability.

1.....2.....3.....4.....5.....6.....7

21. I have experienced visions for which there were no normal causes.

1.....2.....3.....4.....5.....6.....7

22. At least once, I have experienced my 'self' being separate from my body.

1.....2.....3.....4.....5.....6.....7

23. I could demonstrate ESP under controlled conditions.

1.....2.....3.....4.....5.....6.....7

24. I could demonstrate PK under controlled conditions.

1.....2.....3.....4.....5.....6.....7

25. I have experienced telepathy with another person.

1.....2.....3.....4.....5.....6.....7

26. I have experienced a future event before it happened.

1.....2.....3.....4.....5.....6.....7

27. I have 'seen' or 'heard' the spirits of deceased people, friends and relatives.

1.....2.....3.....4.....5.....6.....7

28. At times, objects that I knew could only have been in one location, went missing and were unaccountably found in another location.

1.....2.....3.....4.....5.....6.....7

29. I am sometimes able to influence people, and objects just by thinking about them.

1.....2.....3.....4.....5.....6.....7

30. ESP is a fact.

1.....2.....3.....4.....5.....6.....7

31. Life after death is a fact.

1.....2.....3.....4.....5.....6.....7

32. Telepathy is possible.

1.....2.....3.....4.....5.....6.....7

33. Precognition is possible.

1.....2.....3.....4.....5.....6.....7

34. Contact with spirits is possible.

1.....2.....3.....4.....5.....6.....7

35. Psychokinesis is a fact.

1.....2.....3.....4.....5.....6.....7

36. I am psychic.

1.....2.....3.....4.....5.....6.....7

37. I possess marked psychokinetic ability.

1.....2.....3.....4.....5.....6.....7

N.B. If you have had paranormal experiences that you feel you need to go into further details about, please feel free to describe your experiences in more detail on separate sheets of paper. Also, in some cases I may be interested in following up on your accounts of paranormal experiences (perhaps in a personal interview at your home).

Tick this box if you would welcome further inquiries into your experiences ☐.

Section 2. (Takes approximately 9 minutes)

In the following section you are requested to indicate whether a statement is true or false with respect to you. Do this by placing a TICK in the box (e.g. ☐) at the end of each statement if that statement was true of you as a child. If the statement does not apply to you as a child then leave the box blank.

---

38. I can remember clearly one or more things that happened to me when I was three years of age or younger. ☐
39. When I remember back to when I was six or seven years of age, I can "see" and "hear" again what I saw and heard then, and can feel again the feelings (happy, sad, afraid, etc.) I felt then. ☐
40. Although I have grown, I still feel almost the same as I did when I was younger. ☐
41. When I was younger, I believed in such things as elves, witches, leprechauns, fairies, etc. ☐
42. When I was younger, I would dream or make believe I was flying so clearly that *I felt as if I really did fly.* ☐
43. When I was younger, I enjoyed fairy tales. ☐
44. When I was younger, I was *very good* at make-believe and imagining. ☐
45. When I was younger, I spent time thinking about such things as the meaning of life and death. ☐
46. When I was a young child (below age 8) I liked playing make-believe games such as cowboys, school teacher, husband/wife, etc. I liked them better than games without make-believe such as draughts, building things, ball games, hopscotch, etc. ☐
47. When I was playing make-believe games as a young child, I would make-believe so well that what I pretended seemed real to me. ☐
48. When I was younger, I lived in a make-believe world *much or most* of the time. ☐
49. When I was *much younger*, I believed that my doll(s) or stuffed animal(s) were alive and had feelings (that is, they could feel pain, lonely, happy, etc.). ☐
50. When I was younger, I had a pretend friend or companion (such as a make-believe person, animal, or object) which I talked to, or took along with me. ☐

51. When I was younger, I would at times pretend and in some ways believe I was someone else such as a prince, princess, Snow White, Peter Pan, etc. ☐
52. When I was younger, I would have liked to take ballet lessons or I do or did *enjoy* taking ballet lessons. ☐
53. When I was younger, I would spend at least half of the time I was awake pretending or imagining. ☐
54. At times, when I was younger, it was hard for me to tell if something had actually happened or if I had *imagined* it happened. ☐

N.B. The remaining statements in this section require responses to be given in the format found in section 1. i.e. please circle the number on the scale that best represents your degree of agreement with that statement. The full scale is given again below as a reminder.

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Slightly Disagree	Uncertain	Slightly Agree	Moderately Agree	Strongly Agree

---

55. I am a logical person.

1.....2.....3.....4.....5.....6.....7

56. I rarely use harsh language.

1.....2.....3.....4.....5.....6.....7

57. I can make decisions easily.

1.....2.....3.....4.....5.....6.....7

58. I am generally aware of others feelings.

1.....2.....3.....4.....5.....6.....7

59. I am never conceited about my appearance.

1.....2.....3.....4.....5.....6.....7

60. I have a strong need for security.

1.....2.....3.....4.....5.....6.....7

61. I can be quite aggressive at times.

1.....2.....3.....4.....5.....6.....7

62. I can easily express tender feelings.

1.....2.....3.....4.....5.....6.....7

63. My feelings are not easily hurt.

1.....2.....3.....4.....5.....6.....7

64. I am very neat and tidy.

1.....2.....3.....4.....5.....6.....7

---

65. I am easily flattered.

1.....2.....3.....4.....5.....6.....7

66. I am a forceful person.

1.....2.....3.....4.....5.....6.....7

67. I am very sympathetic to the needs of others.

1.....2.....3.....4.....5.....6.....7

68. On most important issues, I am willing to take a stand.

1.....2.....3.....4.....5.....6.....7

69. People often say I have a gentle, almost childlike manner.

1.....2.....3.....4.....5.....6.....7

70. I can be very affectionate.

1.....2.....3.....4.....5.....6.....7

71. In most things I like to take the lead.

1.....2.....3.....4.....5.....6.....7

72. I am really quite ambitious.

1.....2.....3.....4.....5.....6.....7

What follows next is a series of things that might happen to you now and again. I want you to say whether you regard those events as good (you'd like them to happen more often and would seek to encourage them) or bad (you'd rather not experience them and would try to avoid them). Once again, you must indicate your attitude on a seven point scale by <u>circling the number</u> that best represents your feelings; <u>N.B. the scale has changed and is given below for your convenience!</u>

1	2	3	4	5	6	7
Very Bad	Moderately Bad	Slightly Bad	Indifferent	Slightly Good	Moderately Good	Very Good

73. A smell or scent makes you vividly recall an old memory.

1.....2.....3.....4.....5.....6.....7

74. A thought or image just pops into your head spontaneously.

1.....2.....3.....4.....5.....6.....7

75. You are trying to solve a puzzle and find that it can only be solved using visual imagery- simply thinking analytically or logically about it won't work.

1.....2.....3.....4.....5.....6.....7

76. You try to imagine that your body is really heavy so well that you are sure you are paralysed!

1.....2.....3.....4.....5.....6.....7

77. You start to daydream while putting 1000 stamps on 1000 envelopes, and find much later that you've finished, but you'd forgotten that you were doing such a boring task as your daydream was so vivid.

1.....2.....3.....4.....5.....6.....7

78. You fantasise about something happening so well that it is almost like watching a gripping movie.

1.....2.....3.....4.....5.....6.....7

79. You think back to an event in your childhood and the memory is so vivid it is like actually being a child again.

1.....2.....3.....4.....5.....6.....7

80. You stare at a picture and then look away from it, but it still persists in your mind's eye.

1.....2.....3.....4.....5.....6.....7

Section 3. (Takes approx. 9 minutes)

In this section you will be asked questions of a personal nature. Be assured, information you provide is administered in the strictest confidence by myself alone, and your name and address will not be identified with this questionnaire (they are held separately and identified with you only by an arbitrary I.D. number). You may wish to fill this section in privately - to ensure that you give the most honest response free from the observation of others. If you have further queries regarding this section please feel free to contact me on 031 650 3063, and I will return your call at the earliest moment. Thank you.

The items in the following section describe a number of events which may disrupt a child's life and affect their happiness and functioning. Using the response scale shown below please indicate how many times (if at all) each event occurred in your childhood or adolescence. Simply circle the appropriate response in the response column to the right of each item.

Response scale

Number of occurrences: (A) none
(B) one
(C) two - five
(D) six - ten
(E) more than ten

81. Moved to a new town, started in a new school, had to make new friends.	A B C D E
82. Your parents were divorced or separated.	A B C D E
83. A *serious illness* or *operation* which required you to be hospitalised or miss much school.	A B C D E
84. Had a major injury or accident.	A B C D E
85. Had an abortion or miscarriage.	A B C D E
86. A parent, sibling or other close family member was hospitalised.	A B C D E
87. Major injury or accident of parent, sibling or other close family member.	A B C D E
88. The death of a parent, sibling or other close personal friend.	A B C D E
89. Major injury or accident of a close personal friend.	A B C D E
90. Serious illness or operation of a close personal friend.	A B C D E
91. The death of a close personal friend.	A B C D E
92. Having your house destroyed or damaged (e.g. by flood, fire or earthquake).	A B C D E
93. Having your home robbed or vandalised.	A B C D E

94. Being confronted by a person and robbed of money or belongings. A B C D E

95. Being struck or beaten up by someone other than a family member (e.g. school bully, stranger). A B C D E

96. Being spanked by your parents or other caregivers (e.g. older sibling, baby-sitter, etc.) severely enough to cause bruises or other injuries. A B C D E

97. Being struck or hit by your parents, other family members or caregivers. A B C D E

98. Being struck or hit by someone who wanted to injure you and requiring medical attention (broken bone, stitches, etc.). A B C D E

99. Being yelled and screamed at by a parent, other family member or caregiver. A B C D E

100. Observed or hearing your parents fighting. A B C D E

101. Seeing one of your parents being physically abused by the other (struck, shoved, etc.) A B C D E

102. Seeing one of your siblings being physically abused or abused by your parent(s). A B C D E

103. A family member or relative exposed their sex organs to you when you did not want or expect this to happen. A B C D E

104. Another person exposed their sex organs to you when did not want or expect this to happen. A B C D E

105. A family member or relative touched you in a sexual way when you did not want or expect this to happen. A B C D E

106. Another person touched you in a sexual way when you did not want or expect this to happen. A B C D E

107. A family member or relative made you or asked you to engage in a sexual activity when you did not want to. A B C D E

108. Another person made you or asked you to engage in a sexual activity when you did not want to. A B C D E

109. You were isolated from friends or playmates for a long period(s) of time (e.g. due to quarantine, chronic illness, foreign travel, living on an isolated farm, etc.) A B C D E

110. Continuation of item 109: Average period of time spent in isolation was;
(A) no isolation; (B) 1 week; (C) 1 month;
(D) 6 months; (E) more than 1 year. A B C D E

N.B. Please complete the following section using the true-false tick method. i.e. leave the box blank if the statement does not apply to you- tick ONLY if true of you.

---

111. Before voting I thoroughly investigate the qualifications of all candidates. ☐
112. It is sometimes hard for me to go on with my work if I am not encouraged. ☐
113. On occasions I have doubts about my ability to succeed in life. ☐
114. I am always careful about my manner of dress. ☐
115. If I could get into a movie without paying and be sure I was not seen I would do it. ☐
116. I like to gossip at times. ☐
117. No matter who I'm talking to, I'm a good listener. ☐
118. There have been occasions when I have taken advantage of someone else. ☐
119. I always try to practice what I preach. ☐
120. I sometimes try to get even rather than forgive and forget. ☐
121. I am always courteous, even to people who are disagreeable. ☐
122. There have been occasions when I have felt like smashing things. ☐
123. I never resent being asked to return favour. ☐

To answer the next question you must CIRCLE one of the seven possible responses to indicate your level of life satisfaction.

124. How do you feel about your life as a whole?
1. ....Terrible
2. ....Very dissatisfied
3. .... Mostly dissatisfied
4. .... Mixed feelings- equally satified and dissatisfied
5. .... Mostly satisfied
6. .... Very satisfied
7. .... Delighted

Section 4. (Takes approximately 9 minutes)

In the following section you will be given a series of statements that may or not be true of you. In filling in this section use the standard agree-disagree scale by circling the number that best represents your degree of agreement with that statement. The scale is given below for your convenience;

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Slightly Disagree	Uncertain	Slightly Agree	Moderately Agree	Strongly Agree

---

125. I can usually acheive what I want when I work hard for it.

1.....2.....3.....4.....5.....6.....7

126. Once I make plans I am almost certain to make them work.

1.....2.....3.....4.....5.....6.....7

127. I prefer games involving some luck over games of pure skill.

1.....2.....3.....4.....5.....6.....7

128. I can learn almost anything if I set my mind to it.

1.....2.....3.....4.....5.....6.....7

129. My major accomplishments are entirely due to my hard work and ability.

1.....2.....3.....4.....5.....6.....7

130. I usually do not set goals because I have a hard time following through on them.

1.....2.....3.....4.....5.....6.....7

131. Bad luck has sometimes prevented me from achieving things.

1.....2.....3.....4.....5.....6.....7

132. Almost anything is really possible for me if I really want it.

1.....2.....3.....4.....5.....6.....7

133. Most of what will happen in my career is beyond my control.

1.....2.....3.....4.....5.....6.....7

134. I find it pointless to keep working on something that is too difficult for me.

1.....2.....3.....4.....5.....6.....7

135. In my personal relationships, the other person usually has more control over the relationship than I do.

1.....2.....3.....4.....5.....6.....7

136. I have no trouble making and keeping friends.

1.....2.....3.....4.....5.....6.....7

137. I'm not good at guiding the course of a conversation with several others.

1.....2.....3.....4.....5.....6.....7

138. I can usually develop a close personal relationship with someone I find appealing.

1.....2.....3.....4.....5.....6.....7

139. I can usually steer a conversation towards the topics I want to talk about.

1.....2.....3.....4.....5.....6.....7

140. When I need assistance in something, I often find it difficult to get others to help.

1.....2.....3.....4.....5.....6.....7

141. If there is someone I want to meet I can usually arrange it.

1.....2.....3.....4.....5.....6.....7

142. I often find it hard to get my point of view across to others.

1.....2.....3.....4.....5.....6.....7

143. In attempting to smooth over a disagreement I can sometimes make it worse.

1.....2.....3.....4.....5.....6.....7

144. I find it easy to play an important part in most group situations.

1.....2.....3.....4.....5.....6.....7

145. By taking an active part in political and social affairs we, the people, can control world events.

1.....2.....3.....4.....5.....6.....7

146. The average citizen can have an influence on government decisions.

1.....2.....3.....4.....5.....6.....7

147. It is difficult for us to have much control over the things politicians do in office.

1.....2.....3.....4.....5.....6.....7

148. Bad economic conditions are caused by world events that are beyond our control.

1.....2.....3.....4.....5.....6.....7

149. With enough effort we can wipe out political corruption.

1.....2.....3.....4.....5.....6.....7

150. One of the major reasons we have wars is because people don't take enough interest in politics.

1.....2.....3.....4.....5.....6.....7

151. There is nothing we, as consumers, can do to keep the cost of living going higher.

1.....2.....3.....4.....5.....6.....7

152. It is impossible to have any real influence over what big businesses do.

1.....2.....3.....4.....5.....6.....7

153. I prefer to concentrate my energy on other things rather than on solving the world's problems.

1.....2.....3.....4.....5.....6.....7

154. In the long run we, the voters, are responsible for bad government on a national as well as local level.

1.....2.....3.....4.....5.....6.....7

155. I am usually completely bored.

1.....2.....3.....4.....5.....6.....7

156. Life seems always exciting to me.

1.....2.....3.....4.....5.....6.....7

157. I have no goals in life at all.

1.....2.....3.....4.....5.....6.....7

158. My personal existence is utterly meaningless and without purpose.

1.....2.....3.....4.....5.....6.....7

159. Every day is constantly new and different.

1.....2.....3.....4.....5.....6.....7

160. If I could choose I would prefer never to have been born.

1.....2.....3.....4.....5.....6.....7

161. After retiring, I would do some of the exciting things I always wanted to do.

1.....2.....3.....4.....5.....6.....7

162. In achieving life goals I have made no progress whatsoever.

1.....2.....3.....4.....5.....6.....7

163. My life is emtpy, filled only with despair.

1.....2.....3.....4.....5.....6.....7

164. If I should die today, I would feel that my life has been very worthwile.

1.....2.....3.....4.....5.....6.....7

165. In thinking of my life I often wonder why I exist.

1.....2.....3.....4.....5.....6.....7

166. As I view the world in relation to my life, the world completely confuses me.

1.....2.....3.....4.....5.....6.....7

167. I am a very responsible person.

1.....2.....3.....4.....5.....6.....7

168. Concerning peoples freedom to make their own choices, I believe people have free will.

1.....2.....3.....4.....5.....6.....7

169. With regard to death, I am prepared and not afraid to die.

1.....2.....3.....4.....5.....6.....7

170. With regard to suicide, I have thought of it seriously as a way out of life's problems.

1.....2.....3.....4.....5.....6.....7

171. I have a great ability to find meaning, purpose, or mission in life.

1.....2.....3.....4.....5.....6.....7

172. My life is in my hands and I am in control of it.

1.....2.....3.....4.....5.....6.....7

173. Facing my daily tasks is a source of pleasure and satisfaction.

1.....2.....3.....4.....5.....6.....7

174. I have discovered no mission or purpose in my life.

1.....2.....3.....4.....5.....6.....7

Section 5. (Takes approx. 9 minutes)

In this last section all questions require a response using the familiar 7 point scale. Please circle a number for each statement, according to the following scale;

1	2	3	4	5	6	7
Strongly Disagree	Moderately Disagree	Slightly Disagree	Uncertain	Slightly Agree	Moderately Agree	Strongly Agree

---

175. As a child my parents use to 'go along' with my fantasies.

1.....2.....3.....4.....5.....6.....7

176. My parents discouraged my attempts at make-believe when I was young.

1.....2.....3.....4.....5.....6.....7

177. My parents emphasized concrete, analytically thinking over more creative fantastic 'thought'.

1.....2.....3.....4.....5.....6.....7

178. I can remember my parents showing indifference at my attempts to engage them in my pretend games.

1.....2.....3.....4.....5.....6.....7

179. My parents use to laugh and giggle when I played 'mums and dads', often joining in.

1.....2.....3.....4.....5.....6.....7

180. My parents used to play the old trick of food-as-aeroplane 'coming in to land' in my mouth.

1.....2.....3.....4.....5.....6.....7

181. If I pretended to shoot my father, playing cowboys and indians, he would usually drop down 'dead'.

1.....2.....3.....4.....5.....6.....7

182. If I 'set up shop' as a youngster mum would often come in and buy make-believe 'goods' from me.

1.....2.....3.....4.....5.....6.....7

183. I would often hear my parents telling the neighbours about my imaginary exploits and fantasies.

1.....2.....3.....4.....5.....6.....7

184. My parents often read bedtime stories, vividly telling us of all manner of children's heroes and villains.

1.....2.....3.....4.....5.....6.....7

185. I think that if I had had an imaginary friend as a child, my parents would have been a little worried.

1.....2.....3.....4.....5.....6.....7

In what follows you will find a series of sentences or words that may describe how you have felt over the past month. All you have to do is indicate, using the scale provided, how OFTEN you have felt those things (over the past month). There are five choices for how often you have felt each feeling or sentiment;

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

Over the past MONTH, how OFTEN have you experienced the feeling that...

186. My life is on the right track.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

187. I seem to be left alone when I don't want to be.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

188. I feel I can do whatever I want to.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

189. I think clearly and creatively.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

190. I feel like a failure.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

191. Nothing seems very much fun any more.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

192. I like myself.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

193. I can't be bothered doing anything.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

194. I feel close to people around me.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

195. I feel as though the best years of my life are over.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

196. My future looks good.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

197. I have lost interest in other people and don't care about them.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

198. I have energy to spare.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

199. I smile and laugh a lot.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

200. I wish I could change some parts of my life.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

201. My thoughts go around in useless circles.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

202. I can handle any problems that come up.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

203. My life seems stuck in a rut.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

204. I feel loved and trusted.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

205. I feel there must be something wrong with me.

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

Over the past MONTH how OFTEN have you felt that...

206. satisfied

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

207. lonely

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

208. free-and-easy

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

209. clear-headed

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

210. helpless

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

211. impatient

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

212. useful

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

213. depressed

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

214. loving

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

215. hopeless

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

216. optimistic

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

217. withdrawn

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

218. enthusiastic

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

219. good-natured

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

220. discontented

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

221. confused

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

222. confident

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

223. tense

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

224. understood

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

225. insignificant

Not at all ☐ Occasionally ☐ Some of the time ☐ Often ☐ All of the time ☐

Thank you for your generous assistance. You have been one of the participants in a study which will better enable parapsychologists to understand the causes and consequences of belief in the paranormal.

If you wish to recieve details of the results of this study, **including a brief run down of your own results**, please tick ☐.

Thank you once again.

Tony Lawrence, B.Sc.
Department of Psychology
University of Edinburgh
7 George Square
Edinburgh

EH 8 9JZ Tel. 031 650 3063.

Appendix III

Correlation matrices and standard deviations for the modelling analyses.

Confirmatory Factor Analyses (Chapter 5)

Part I - Correlations and standard deviations for 26 PBS items (five-point scale) N=131

	TOB1	TOB2	TOB3	TOB4	TOB5	TOB6
TOB1	1.000					
TOB2	0.305	1.000				
TOB3	0.417	0.481	1.000			
TOB4	0.044	0.297	0.212	1.000		
TOB5	0.420	0.564	0.442	0.145	1.000	
TOB6	0.093	0.215	0.231	0.122	0.389	1.000
TOB7	0.128	0.256	0.131	0.340	0.183	-0.015
TOB8	0.465	0.154	0.461	0.168	0.158	0.035
TOB9	0.360	0.708	0.519	0.189	0.591	0.323
TOB10	0.367	0.425	0.625	0.103	0.294	0.248
TOB11	0.133	0.264	0.226	0.694	0.235	0.172
TOB12	0.453	0.405	0.382	0.159	0.726	0.318
TOB13	0.080	0.171	0.215	0.158	0.181	0.461
TOB14	0.182	0.304	0.182	0.333	0.307	0.096
TOB15	0.603	0.242	0.281	0.124	0.231	-0.000

TOB16	0.260	0.753	0.475	0.225	0.600	0.240
TOB17	0.295	0.441	0.657	0.272	0.432	0.269
TOB18	0.062	0.242	0.194	0.632	0.162	0.103
TOB19	0.343	0.323	0.227	0.189	0.471	0.307
TOB20	-0.094	0.060	-0.109	-0.066	0.215	0.299
TOB21	0.349	0.512	0.421	0.255	0.540	0.148
TOB22	0.434	0.144	0.337	0.159	0.082	-0.018
TOB23	0.296	0.501	0.266	0.164	0.432	0.062
TOB24	0.283	0.372	0.604	0.196	0.370	0.298
TOB25	0.295	0.544	0.424	0.237	0.507	0.251
TOB26	0.330	0.409	0.396	0.205	0.504	0.137

	TOB7	TOB8	TOB9	TOB10	TOB11	TOB12
TOB7	1.000					
TOB8	0.145	1.000				
TOB9	0.113	0.224	1.000			
TOB10	0.005	0.268	0.482	1.000		
TOB11	0.453	0.156	0.180	0.170	1.000	
TOB12	0.352	0.255	0.455	0.289	0.286	1.000
TOB13	0.063	-0.011	0.209	0.198	0.332	0.212
TOB14	0.801	0.105	0.182	0.053	0.485	0.444
TOB15	0.137	0.546	0.276	0.270	0.141	0.285

TOB16	0.216	0.166	0.787	0.428	0.245	0.517
TOB17	0.172	0.279	0.422	0.568	0.256	0.386
TOB18	0.428	0.133	0.122	0.073	0.811	0.193
TOB19	0.177	0.050	0.395	0.218	0.237	0.525
TOB20	-0.006	-0.119	0.202	-0.059	0.035	0.258
TOB21	0.431	0.289	0.500	0.375	0.360	0.599
TOB22	0.136	0.798	0.178	0.206	0.244	0.201
TOB23	0.298	0.232	0.466	0.162	0.202	0.404
TOB24	0.032	0.300	0.458	0.717	0.233	0.363
TOB25	0.346	0.222	0.529	0.320	0.270	0.520
TOB26	0.329	0.175	0.396	0.309	0.348	0.513

	TOB13	TOB14	TOB15	TOB16	TOB17	TOB18
TOB13	1.000					
TOB14	0.124	1.000				
TOB15	0.136	0.079	1.000			
TOB16	0.187	0.281	0.174	1.000		
TOB17	0.169	0.276	0.150	0.452	1.000	
TOB18	0.259	0.416	0.100	0.229	0.253	1.000
TOB19	0.126	0.197	0.108	0.436	0.119	0.188
TOB20	0.239	0.047	-0.090	0.227	-0.042	0.072
TOB21	0.137	0.469	0.292	0.554	0.361	0.364

TOB22	0.046	0.124	0.659	0.114	0.218	0.167
TOB23	0.053	0.231	0.291	0.490	0.157	0.195
TOB24	0.196	0.150	0.250	0.423	0.596	0.195
TOB25	0.112	0.341	0.146	0.534	0.451	0.257
TOB26	0.163	0.385	0.182	0.454	0.356	0.331

	TOB19	TOB20	TOB21	TOB22	TOB23	TOB24
TOB19	1.000					
TOB20	0.362	1.000				
TOB21	0.446	0.053	1.000			
TOB22	0.048	-0.103	0.252	1.000		
TOB23	0.267	0.048	0.452	0.201	1.000	
TOB24	0.172	0.019	0.423	0.222	0.132	1.000
TOB25	0.413	0.245	0.478	0.117	0.329	0.361
TOB26	0.311	-0.062	0.669	0.137	0.385	0.392

	TOB25	TOB26
TOB25	1.000	
TOB26	0.469	1.000

Standard Deviations:

TOB1 = 1.294, TOB2 = 1.097, TOB3 = 1.220, TOB4 = 0.8, TOB5 = 1.188, TOB6 = 0.957, TOB7 = 1.027, TOB8 = 1.424, TOB9 = 1.180, TOB10 = 1.336, TOB11 = 1.020, TOB12 = 1.201, TOB13 = 1.104, TOB 14 = 0.980, TOB15 = 1.432, TOB16 = 1.159, TOB17 = 1.038, TOB18 =1.014, TOB19 =

1.199, TOB20 = 1.032, TOB21 = 1.093, TOB22 = 1.406, TOB23 = 1.005, TOB24 = 1.080, TOB25 = 1.196, TOB26 = 1.010

Part II - Correlation matrix and standard deviations for 26 item PBS (seven point scale)

N = 225

	q1	q2	q3	q4	q5	q6
q1	1.000					
q2	0.306	1.000				
q3	0.445	0.506	1.000			
q4	0.134	0.264	0.267	1.000		
q5	0.251	0.515	0.333	0.229	1.000	
q6	0.153	0.346	0.252	0.337	0.367	1.000
q7	0.201	0.387	0.271	0.372	0.319	0.303
q8	0.330	0.296	0.550	0.134	0.111	0.195
q9	0.220	0.715	0.416	0.248	0.543	0.367
q10	0.295	0.378	0.571	0.196	0.322	0.297
q11	0.116	0.211	0.145	0.608	0.201	0.327
q12	0.395	0.452	0.361	0.155	0.578	0.313
q13	0.020	0.121	0.069	0.173	0.213	0.545
q14	0.243	0.339	0.205	0.313	0.274	0.289
q15	0.391	0.095	0.224	0.074	-0.053	-0.049
q16	0.217	0.687	0.336	0.216	0.623	0.361
q17	0.274	0.309	0.526	0.229	0.288	0.239
q18	0.123	0.272	0.226	0.448	0.146	0.245
q19	0.389	0.300	0.202	0.174	0.330	0.158
q20	0.060	0.237	0.074	-0.070	0.271	0.245
q21	0.204	0.346	0.277	0.338	0.284	0.229
q22	0.313	0.163	0.329	0.134	0.034	0.050
q23	-0.244	-0.212	-0.188	0.034	-0.143	-0.075
q24	0.287	0.309	0.619	0.223	0.298	0.270
q25	0.346	0.447	0.447	0.183	0.364	0.251
q26	0.332	0.408	0.486	0.235	0.358	0.269

	q7	q8	q9	q10	q11	q12
q7	1.000					
q8	0.161	1.000				
q9	0.357	0.209	1.000			
q10	0.238	0.395	0.328	1.000		
q11	0.381	0.147	0.234	0.183	1.000	

q12	0.352	0.296	0.444	0.337	0.154	1.000
q13	0.117	0.030	0.194	0.147	0.268	0.163
q14	0.660	0.065	0.347	0.160	0.386	0.243
q15	0.117	0.436	0.053	-0.009	0.070	0.088
q16	0.333	0.117	0.715	0.296	0.209	0.476
q17	0.364	0.376	0.279	0.494	0.199	0.358
q18	0.287	0.166	0.289	0.254	0.631	0.101
q19	0.268	0.123	0.297	0.208	0.225	0.269
q20	0.014	-0.060	0.270	0.106	0.034	0.214
q21	0.381	0.157	0.427	0.209	0.353	0.305
q22	0.186	0.628	0.080	0.136	0.190	0.137
q23	0.079	-0.047	-0.133	-0.087	0.016	-0.317
q24	0.283	0.388	0.352	0.619	0.098	0.345
q25	0.373	0.170	0.472	0.328	0.217	0.315
q26	0.376	0.286	0.445	0.331	0.233	0.418

	q13	q14	q15	q16	q17	q18
q13	1.000					
q14	0.144	1.000				
q15	-0.020	0.000	1.000			
q16	0.247	0.357	0.004	1.000		
q17	0.118	0.255	0.080	0.296	1.000	
q18	0.261	0.399	0.116	0.227	0.234	1.000
q19	0.068	0.256	0.108	0.308	0.118	0.211
q20	0.205	0.205	-0.170	0.273	0.131	0.005
q21	0.213	0.335	0.084	0.320	0.146	0.292
q22	-0.032	0.105	0.626	-0.019	0.137	0.190
q23	-0.110	0.041	-0.030	-0.225	-0.156	0.091
q24	0.123	0.169	0.128	0.207	0.428	0.167
q25	0.160	0.289	0.028	0.365	0.237	0.250
q26	0.152	0.313	0.160	0.355	0.310	0.245

	q19	q20	q21	q22	q23	q24
q19	1.000					
q20	0.195	1.000				
q21	0.363	0.212	1.000			
q22	0.097	-0.107	0.156	1.000		

q23	-0.107	-0.185	-0.123	-0.009	1.000	
q24	0.125	0.085	0.290	0.224	-0.050	1.000
q25	0.454	0.206	0.495	0.089	-0.145	0.362
q26	0.355	0.181	0.567	0.151	-0.240	0.394

	q25	q26
q25	1.000	
q26	0.429	1.000

Standard Deviations:

V1 = 1.585, V2 = 1.574, V = 1.410, V5 = 1.575, V6 = 1.624, V7 = 1.609, V8 = 1.880, V9 = 1.523, V10 = 1.847, V11 = 1.539, V12 = 1.570, V13 = 1.522, V14 = 1.459, V15 = 1.870, V16 = 1.606, V17 = 1.563, V18 = 1.460, V19 = 1.540, V20 = 1.338, V21 = 1.501, V22 = 1.713, V23 = 1.388, V24 = 1.558, V25 = 1.578, V26 = 1.470

First Edinburgh Student Study (Chapter 6) N= 80

	Trauma	Fantasy	Experience
Fantasy	.356	---	---
Experience	.343	.315	---
Belief	.215	.271	.534

Standard Deviations: Trauma = 8.13, Fantasy = 6.46, PB= 5.75, PE= 9.55

Society for Psychical Research Survey (from chapter 7) N=58

	Experience	EPBS	ASGS-B	STCE
EPBS	0.551	---	---	---
ASGS-B	0.316	0.373	---	---
STCE	0.440	0.168	-0.051	---
Fantasy	0.272	0.271	0.180	0.131

Standard deviations:

PES = 18.027, EPBS = 7.455, ASGS-B = 7.291, CF = 3.154, STCE = 10.209

Edinburgh Survey (Chapter 8)

Part 1 - Irwin vs. Lawrence (Basic Models) N=129.

	PES	EPBS	ASGS-B	CF	STCE
PES	1.000				
EPBS	0.723	1.000			
ASGS-B	0.508	0.676	1.000		
CF	0.476	0.355	0.245	1.000	
STCE	0.250	0.273	0.175	0.274	1.000

Standard deviations:

PES = 16.742, EPBS = 9.602, ASGS-B = 8.767, CF = 3.315, STCE = 7.305

Chapter 8 Part II - Specified Trauma Model Analyses (N=129).

	PES	EPBS	ASGS-B	CF	FAMPHYS	LOSSFRI	LOSSFAM	ISOLAT	ILLACC	PARDIV	LOSSHOM	ROBBER
PES	1.000											
EPBS	0.723	1.000										
ASGS-B	0.508	0.676	1.000									
CF	0.476	0.355	0.245	1.000								
FAMPHYS	0.184	0.227	0.161	0.209	1.000							
LOSSFRI	0.125	0.099	-0.044	0.018	-0.005	1.000						
LOSSFAM	0.019	0.037	0.049	-0.074	0.044	0.452	1.000					
ISOLAT	0.196	0.170	0.124	0.223	0.100	0.302	0.249	1.000				
ILLACC	0.064	0.068	0.051	0.102	0.048	0.051	0.156	0.298	1.000			
PARDIV	0.134	0.117	0.101	0.281	0.163	0.181	-0.010	0.039	-0.093	1.000		
LOSSHOM	0.143	0.024	-0.014	0.051	0.099	0.138	0.042	0.006	0.040	0.223	1.000	
ROBBER	0.114	0.142	0.046	0.083	0.139	-0.049	0.043	0.010	-0.026	0.063	0.145	1.000

Standard deviations:

PES = 16.742, EPBS = 9.602, ASGS-B = 8.767, CF = 3.315, FAMPHYS = 4.401, LOSSFRI = 1.288, LOSSFAM = 1.656, ISOLAT = 2.469, ILLACC = 0.650, PARDIV = 0.649, LOSSHOM = 1.121, ROBBER = 0.317.

Chapter 9 Correlations and standard deviations for all analyses

Part I - Whole sample Personal Consequences Model (N=129)

	pes	epbs	asgs-b	life 3	soc-p	soc-i
pes	1.000					
epbs	0.723	1.000				
asgs-b	0.508	0.676	1.000			
life 3	0.020	0.034	0.057	1.000		
soc-p	0.005	-0.061	-0.006	0.289	1.000	
soc-i	0.176	0.076	0.028	0.301	0.351	1.000
pil	0.066	-0.017	0.032	0.553	0.426	0.483
wellbe	-0.031	-0.000	-0.003	0.604	0.293	0.484

	pil	wellbe
pil	1.000	
wellbe	0.624	1.000

Standard Deviations:

PES = 16.742, EPBS = 9.602, ASGS-B = 8.767, LIFE3 = 0.906, SOC-P = 7.341, SOC-I = 8.763, PIL = 14.634, WELLBE = 22.353

Correlations and S.D.'s for PCM Model for Females (N=71)

	pes	epbs	asgs-b	life 3	soc-p	soc-i
pes	1.000					
epbs	0.662	1.000				
asgs-b	0.523	0.663	1.000			
life 3	0.116	0.169	0.161	1.000		
soc-p	0.165	0.058	0.179	0.307	1.000	
soc-i	0.300	0.178	0.093	0.354	0.357	1.000
pil	0.203	0.111	0.147	0.573	0.447	0.599
wellbe	0.101	0.198	0.190	0.643	0.370	0.560

	pil	wellbe
pil	1.000	

wellbe	0.679	1.000

Standard Deviations:
PES = 16.806, EPBS = 9.289, ASGS-B = 7.715, LIFE3 = 0.971, PIL = 15.444, WELLBE = 25.966, SOC-P = 6.903, SOC-I = 9.848.

Correlations and S.D.'s for PCM Model for Males (N=45)

	pes	epbs	asgs-b	life 3	soc-p	soc-i
pes	1.000					
epbs	0.791	1.000				
asgs-b	0.508	0.714	1.000			
life 3	-0.090	-0.157	0.023	1.000		
soc-p	-0.158	-0.126	-0.113	0.225	1.000	
soc-i	-0.009	-0.048	0.058	0.102	0.348	1.000
pil	-0.015	-0.115	-0.046	0.470	0.345	0.258
wellbe	-0.249	-0.301	-0.198	0.528	0.163	0.212

	pil	wellbe
pil	1.000	
wellbe	0.582	1.000

Standard Deviations:
PES = 15.441, EPBS = 9.614, ASGS-B = 10.526, LIFE3 = 0.809, PIL = 12.381, WELLBE = 15.248, SOC-P = 7.921, SOC-I = 6.918

Appendix IV

The Tobacyk (1988) Revised Paranormal Belief Scale:
N.B. The item numbers are exactly the same as the variable numbers ("v's") given in figures 1, 2, and 3.

All items are rated on a seven point scale thus:
1. Strongly Disagree
2. Moderately Disagree.
3. Disagree
4. Uncertain
5. Agree.
6. Moderately Disagree
7. Strongly Disagree

****

1. The soul continues to exist though the body may die.
2. Some individuals are able to levitate and travel (lift) objects through mental forces.
3. Black magic really exists.
4. Black cats can bring you bad luck.
5. Your mind or soul can leave your body and travel (astral projection).
6. The abominable snowman of Tibet exists.
7. Astrology is a way to accurately predict the future.
8. There is a devil.
9. Psychokinesis, the movement of objects through psychic powers, does exist.
10. Witches do exist.
11. If you break a mirror, you will have bad luck.
12. During altered states, such as sleep or trances, the spirit can leave the body.
13. The Loch Ness monster of Scotland exists.
14. The horoscope accurately tells a person's future.
15. I believe in God.
16. A person's thoughts can influence the movement of a physical object.
17. Through the use of formulas and incantations, it is possible to cast spells on persons.
18. The number "13" is unlucky.
19. Reincarnation does occur.
20. There is life on other planets.
21. Some psychics can accurately predict the future.
22. There is a heaven and a hell.
23. Mind reading is not possible.
24. There are actual cases of witchcraft.
25. It is possible to communicate with the dead.
26. Some people have an unexplained ability to predict the future.

N.B. Item 23 is reversed for scoring.

Appendix V

Fully Published Papers

Pergamon

0191-8869(95)00034-8

Person. individ. Diff. Vol. 19, No. 2, pp. 209–215, 1995

0191-8869/95 $9.50 + 0.00

MODELLING CHILDHOOD CAUSES OF PARANORMAL BELIEF AND EXPERIENCE: CHILDHOOD TRAUMA AND CHILDHOOD FANTASY

Tony Lawrence,* Claire Edwards, Nicholas Barraclough, Sarah Church and Francesca Hetherington

Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, Scotland

(Received 19 September 1994)

Summary—Using covariance structure modelling we sought to test the *childhood factors model* of paranormal belief development outlined by Irwin (*Journal of American Society for Psychical Research, 86*, 199–208; *87*, 1–39). Eighty-two students at the University of Edinburgh were administered three questionnaires relating to childhood trauma, childhood fantasy, and paranormal belief and experience. Eighty suitable questionnaires were completed and analysed using EQS (Bentler, *EQS Structural Relations Program Manual*, 1989). The first test of Irwin's model showed that the model did not provide an adequate fit to our sample data. In particular, EQS suggested *dropping* a direct causal link from fantasy to belief and *adding* a direct link from trauma and paranormal experience. Both modifications were intuitively plausible and were implemented in a *post hoc* modified model. This new model gave an excellent fit to the data. In addition, our study replicated Irwin's (*Journal of the American Society for Psychical Research, 86*, 199–208, 1992) finding of a small but significant correlation between childhood trauma and paranormal belief, and extends previous findings showing a small correlation between childhood fantasy and paranormal belief (and experience) to the realm of childhood fantasy, thus addressing the childhood factors model proper. In conclusion, we offer up our new model for future attempts at replication, and strongly advocate the switch to a model building approach to better understand paranormal belief and experience.

INTRODUCTION

"When it's grey I know all I want to do—resort to make believe."
(From the song *A Love for all Seasons*, by Right Said Fred.)

Any particular person's history of paranormal belief and experience must have some beginning, and some causative factor responsible for that beginning. Until quite recently researchers on the topic of paranormal belief (PB) and experience (PE) remained unaware, or uninterested, in the possible early developmental causes of PB and PE, much of the research in this area reflecting what Irwin (1993) calls the *cognitive deficits hypothesis*—the view, influenced by a broadly sceptical research agenda, that those with high PB or PE are '... illogical, irrational, credulous, uncritical, and foolish' (Irwin, 1993, p. 16). Recently, however, research conducted by Irwin (1992) has shown the potential importance of childhood factors underlying PB, and so it is to a subset of these potentially causative correlates of PB and PE that our paper is addressed—namely, childhood trauma and childhood fantasy.

Research addressing the childhood causes of PB and PE can be traced back to a small study reported by Wilson and Barber (1983). In their study Wilson and Barber (1983) interviewed 27 women whom they had rated as excellent hypnotic *S*s (with a comparison group of 25 women not so rated). The interviews centred around childhood and adult memories, fantasies and psychic experiences. As Wilson and Barber (1983) note, with reference to their excellent hynotic *S*s:

> "Their extensive and deep involvement in fantasy seems to be their basic characteristic and their other major talents—their ability to hallucinate voluntarily, their superb hypnotic performances, their vivid memories of their life experiences, and their talents as psychics or sensitives—seem to derive from or to grow out of their profound fantasy life." (Wilson & Barber, 1983, p. 340).

One of the most significant factors in differentiating between fantasisers and non-fantasisers in Wilson's and Barber's group was the incidence of childhood trauma [previously found by Hilgard

* To whom all correspondence should be addressed.

(1970, 1974)]. Of the 27 fantasisers, nine (33%) reported having experienced some difficult or stressful event in their early life, ranging from physical abuse by parents, or a mother with emotional problems, to unstable living conditions. No *Ss* in the comparison group reported having experienced any childhood trauma [see Lynn and Rhue (1988) for a replication].

Since Wilson and Barber's study, two studies relating fantasy proneness to PB have been conducted by Irwin (1990, 1991). In the first study, Irwin (1990) administered Myers' (1983) version of the Inventory of Childhood Memories and Imaginings (ICMI:C) to 92 people, along with Tobacyk and Milford's (1983) Revised Paranormal Belief Scale (PBS). Irwin found Pearson correlations of 0.48, 0.58, and 0.43 for the Psi Belief, Spiritualism, and Precognition subscales of the PBS respectively with scores on the ICMI:C (all significant at $P < 0.001$), suggesting a definite link between imaginal involvement and the core constructs of PB. In a follow up study, Irwin (1991) administered the ICMI:C and PBS to a sample of 122 students and found a similar pattern of correlations—the PBS Psi Belief, Spiritualism, and Precognition subscales correlated 0.40, 0.42 and 0.34 with the ICMI:C (all $P < 0.001$). Another study by Rao (1992) looked at the relationship between fantasy proneness and PE alone as measured by Neppe's (1983) Subjective Paranormal Experiences Questionnaire (SPEQ). Rao found a correlation of 0.46 ($P < 0.01$) between the SPEQ and the ICMI:C, replicating Wilson and Barber's finding that fantasy prone participants reported more psychic experiences.

On the topic of childhood trauma and PB Irwin (1992) administered the revised PBS (Tobacyk, 1988), and the Survey of Traumatic Childhood Experiences (STCE) (Council & Edwards, 1987) to 72 students. He found a Pearson correlation of 0.21 ($P < 0.05 < 0.08$) between the full scale PBS score (including items of dubious paranormality about the Loch Ness monster and life on other planets) and the STCE; a non-significant result which is at least suggestive. Those items most strongly correlated with PB were the STCE items relating to intrafamilial physical abuse ($r = 0.31$, $P < 0.05$). Thus there is evidence to support the view that childhood trauma has an influence on the subsequent development of PB. In suggesting his own model of these relationships Irwin concludes:

> "... Fantasy proneness ... is a central mediating variable in the development of paranormal beliefs. Fantasy proneness may be kindled in early childhood through parental encouragement and/or the incidence of traumatic experiences. In particular, the latter will instill a strong need for interpersonal control. Given a conducive social context, fantasy proneness and the associated need for control may lead to the formation of paranormal beliefs. Fantasy proneness ... also may be a factor in the occurrence and/or acknowledgement of parapsychological experiences." (Irwin, 1992, p. 206).

Figure 1 presents a reduced version of Irwin's *childhood factors model*, showing only those relations between factors that we have chosen to model in the present study. With respect to his larger model Irwin notes, "Most of the relationships indicated by the arrows have now been documented in the empirical literature, but there remains a need for a single study entailing a path analysis of all variables encompassed by the model." (Irwin, 1992, p. 205).

In this study we attempt a partial confirmation of Irwin's model, using a more advanced technique than path analysis, namely covariance structure modelling (CSM) using EQS (Bentler, 1989), a program specifically designed to implement and test structural equation models using covariances. Using the reduced version of Irwin's model (see Fig. 1) we aim to test the causal structure underlying the relations between childhood trauma, childhood fantasy, PE and PB. No study has to date looked specifically and exclusively at the role of *childhood* fantasy involvement in the development of PB and PE, but we feel that an emphasis on childhood fantasy, as opposed to the mixed effects of child and adult fantasy (as measured by the ICMI) is important given the proximity of childhood fantasy, and its pertinence to traumatic events in childhood.

METHOD

Subjects

The respondents were 82 students and friends of the last four authors, all at the University of Edinburgh. Unfortunately no information on sex or age was taken, but the sample was broadly similar to those that one usually gets from university students.

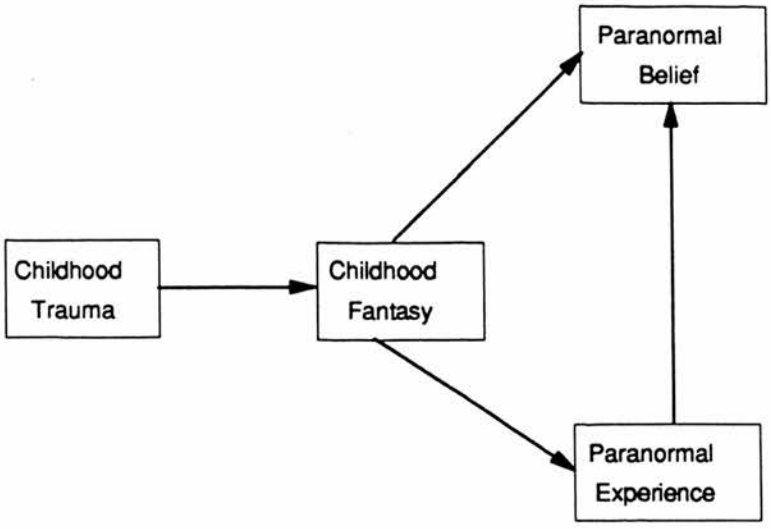

Fig. 1. A simplification of Irwin's (1992) model depicting the childhood causes of paranormal belief.

Materials

A 61-item questionnaire was constructed using items from the following three scales:

(1) Survey of Traumatic Childhood Experiences (STCE). This scale, developed by Council and Edwards (1987), comprises 30 items dealing with events that may cause trauma in a child's life. The full 30 item scale was used in this study. The scale has items relating to intrafamilial sexual and physical abuse, loss related to friends and family, extrafamilial sexual abuse, assault, loss of home and property, isolation and personal malady, adolescent abortion/miscarriage, and parental divorce/separation. Responses are scored on a five point scale of frequency, with (A) none, (B) one, (C) two to five, (D) six to ten, and (E) more than ten traumatic experiences.
(2) Australian Sheep–Goat Scale (ASGS). This scale comprises 18 items referring to three common components of PB; ESP, PK, and Life after Death. Thalbourne and Delin (1993) report reasonable reliability and some evidence showing concurrent validity. Items refer to PB and PE and were marked by a seven point scale.
(3) Childhood fantasy. This construct was measured using only those items from Myers' (1983) ICMI:C that refer explicitly to childhood fantasy. Items are scaled dichotomously, in the true/false format. High scores indicate high fantasy.

Procedure

Participants were given the three-scale questionnaire to fill in, along with a blank envelope for the completed questionnaire to be placed in to preserve the participant's anonymity. Participants were asked to fill in the questionnaire as honestly as possible. Once the questionnaire was completed and sealed in the envelope the participant was briefed about the nature of the study. No participants reported having any particular problem with the content of the scales, though some pointed out that the Australian Sheep–Goat Scale was quite contorted in its wording. Once the data had been collected they were entered in to Statview for calculation of the correlation matrix, before being entered into EQS for covariance structure modelling. At this stage two participants' questionnaires were excluded from the analysis, as they constituted extreme outliers on Trauma and one of the *S*s had (amongst other unlikelies) claimed to have had more than ten abortions/miscarriages! Consequently, we suspected that at least one of the two outliers was not responding honestly. The remaining sample comprises 80 individuals.

Statistical methods

Correlational data from this study were analysed using the technique of covariance structure analysis using EQS (Bentler, 1989). EQS combines factor analysis, multiple regression, and path

Table 1. Pearson correlations between the four main variables

Variable	Trauma	Fantasy	Experience
Fantasy	0.356***	—	—
Experience	0.343***	0.315***	—
Belief	0.215*	0.271*	0.534*****

* $P < 0.06 > 0.05$; ** $P < 0.05$; *** $P < 0.01$; ***** $P < 0.0001$.
Standard deviations: Trauma = 8.13, Fantasy = 6.46, PB = 5.75, PE = 9.55.

analysis techniques, providing a powerful system to test and confirm various causal or structural models. The most important output from a test of any causal model is a list of goodness-of-fit statistics, indicating how well the hypothesized model accounts for the observed correlational structure of the data. When assessing the goodness-of-fit of a causal model there are a number of statistics which we have at our disposal. For the first of these, the χ^2 goodness-of-fit statistic, we require that the χ^2 be *non-significant* for our model to be considered a good fit. Next we have average standardized residual statistics, of which the off diagonal residuals are the most important. These may be treated like residuals in regression analysis, indicating unexplained variance, and are particularly useful in comparing the differences in fit between two or more alternative models. Aikaike's Information Criterion (AIC) χ^2 provides an estimate of goodness-of-fit that takes into account the parsimony of the underlying model (in terms of the number of parameters that must be estimated to arrive at a particular solution). Thus the smaller the value the better. Lastly, Bentler and Bonett's non-normed fit index (NNFI) provides a particularly good estimate of model fit with small sample sizes (appropriate for this sample size). To even be considered as a reasonable fit, the fit index should show values of greater than 0.9.

Lastly, a useful feature of causal modelling is the suggestion of model modifications, whereby EQS suggests changes to the model in order to provide a better fit. Used cautiously, this feature is helpful in achieving models which fit the data well.

RESULTS

From Tables 1 and 2 we can determine how well our study has replicated previous findings on the relation between PE and belief, with fantasy and trauma. In Table 1 we may note that the correlation between fantasy and PB is $r = 0.27$ ($P < 0.05$). This fits reasonably with the studies of Irwin (1990, 1991) showing average correlations of around 0.44 between the ICMI:C and the PBS. The relationship between childhood fantasy and PE was slightly stronger, $r = 0.31$ ($P < 0.01$). This result offers a fair degree of agreement with the findings of Rao (1992) with the ICMI:C ($r = 0.46$). Further, our own reported correlations for childhood fantasy are likely to be smaller simply because we were solely measuring childhood fantasy (the ICMI:C actually has adult fantasy questions too). Obviously we cannot expect memory based estimates of childhood fantasy to correlate as highly with present day PB or PE, as would contemporary fantasy estimates.

With respect to childhood trauma and PB our correlation exactly replicates that of Irwin (1992). This is particularly encouraging, given our use of a much narrower measure of PB. However, we failed to replicate Irwin's *post hoc* finding that the strongest specific traumatic correlate of PB was intrafamilial physical abuse. Our data, presented in Table 2, show that the strongest predictor of PB were the items relating to 'loss of property' (scores for this predictor stemmed mostly from answers to an item relating to the frequency of moving home as a child).

Table 2 gives the descriptive statistics for childhood trauma, split into its various sources. Whilst all of the four main variables gave reasonable approximations to a normal distribution, this was not the case for the specific traumatic variables presented in Table 2. Of these, only intrafamilial physical abuse data showed evidence of being normally distributed, the other variables being more χ^2 in their distribution. Indeed, it is clear that most participants had neither experienced much childhood trauma, nor collectively displayed much variation in trauma experienced. This is particularly the case for familial sexual abuse, almost all participants reporting no sexual abuse in their family as children.

Table 2. Descriptive statistics for specific sources of childhood trauma, with correlations given for paranormal belief and paranormal experience

Source of childhood trauma	Mean	SD	Corrected Spearman ρ (belief)	Corrected Spearman ρ (experience)
Intrafamilial sexual abuse	1.04	0.16	− 0.008	0.058
Intrafamilial physical abuse	2.49	0.92	0.196*	0.194*
Loss (friend)	1.28	0.45	− 0.024	0.23**
Extrafamilial sexual abuse	1.28	0.47	− 0.038	0.086
Loss (family)	1.63	0.56	0.110	0.164
Isolation	1.35	0.67	0.064	0.056
Personal illness	1.40	0.54	− 0.039	0.194*
Parental divorce/miscarriage	1.09	0.19	0.031	0.137
Assault	1.57	0.69	− 0.094	0.18
Loss of property (home moving)	1.52	0.40	0.379****	0.239**
Robbery/mugging	1.12	0.37	0.065	0.251**

* $P < 0.09 > 0.05$; ** $P < 0.05$; **** $P < 0.001$.

Consequently, for correlational analyses on the four main variables we used Pearson product moment correlation coefficients, but for analyses on specific traumatic factors we chose to use Spearman's ρ, corrected for ties.

Lastly, our correlations between the various factors of childhood trauma and PB and experience are interesting in so far as they show that the strength and significance of the correlations reported is much stronger between trauma and experience than for belief. Here we see the inklings of a model relating these factors, and it is to the modelling of these factors proper that we now turn.

Results of causal modelling

First we ran EQS on a slightly simplified version of Irwin's (1992) model, given in Fig 1. The results for this model, in terms of fit statistics, are given in the appropriate column of Table 3. The χ^2 of 5.544 on two degrees of freedom indicates a near significant lack of fit between the model and the observed data. Equally the non-normed fit index is poor; all the worse given that this index is best for reflecting fit at low sample sizes. In short, Irwin's model is a poor fit.

EQS recommended two changes to Irwin's model. Firstly, the Wald test for dropping parts of the model suggested the exclusion of the direct causal link from fantasy to belief. Secondly, the Langrange Multiplier test for adding parts to the model suggested the inclusion of a direct causal link between trauma and PE. Consequently, we made these changes to the model, arriving at a much more streamlined model whereby trauma causes fantasy causes experience causes belief, with a direct link from trauma to experience (we had actually considered modelling this model previously, minus the trauma-experience link).

The results for the test of this model are given in the last column of Table 3. The new model is a considerable improvement. Residuals and fit statistics indicate an excellent degree of fit. AIC statistics, plus residuals, show that this model is considerably better than Irwin's in reproducing the covariance found in the original data. In Fig. 2 we provide a graphical representation of this new model, with path coefficients. It is to a discussion of these results that we now turn.

Table 3. Fit statistics for Irwin's (1992) childhood factors model and the new model

Model fit statistics	Irwin's model (Fig. 1)	New model (Fig. 2)
Average off-diagonal absolute standardized residuals	0.058	0.022
Independence model χ^2 (d.f.)	52.332 (6)	
Independence AIC	40.332	40.332
Model AIC	1.544	− 2.691
χ^2 (d.f.)	5.544 (2)	1.309 (2)
χ^2 P-value	0.0625	0.520
Normal theory RLS χ^2	5.354	1.298
Bentler–Bonett non-normed fit index	0.771	1.045

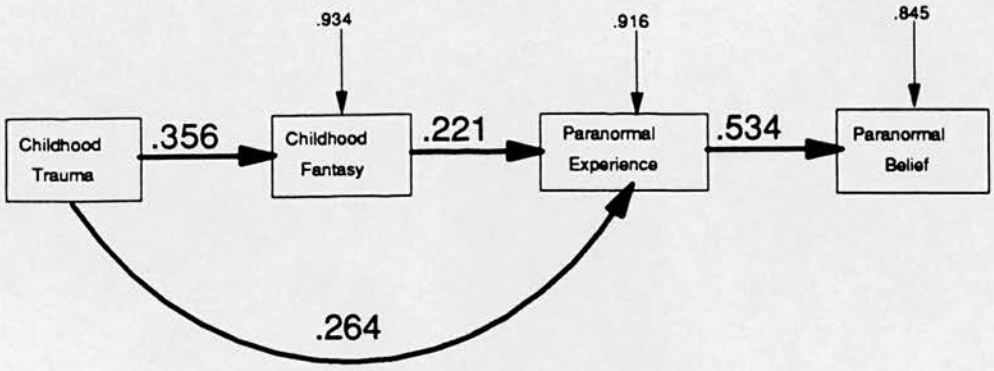

Fig. 2. The new model of the childhood causes of paranormal belief and experience.

DISCUSSION

The correlation between childhood trauma and PB, whilst of marginal significance, was exactly the same as that found by Irwin (1992). Also the previous findings of a relation between fantasy, PB and PE, were confirmed, whilst also being extended to the influence of childhood fantasy

More importantly, Irwin's proposed model of the causes of PB, as it applies to the variables we studied, did not provide a good fit to the data. Our data provided insufficient support for the link between fantasy and PB, and it certainly makes more sense to suppose that any direct effects of heightened fantasy involvement would be felt upon PE, rather than PB. However, of particular interest is the finding that a much better fit stems from the inclusion of a direct link between trauma and PE, not an obvious result by any means. According to the new model trauma has two causal routes in its influence on PE, one direct, the other indirect.

Our finding that trauma gives rise to PE directly, is something which we may explain only once we have been able to determine the specific traumatic causes responsible for the direct route to PE. From a *post hoc* perusal of our data we would like to suggest a possible cause of the two routes to paranormal experience (from trauma). It may be that the direct link stems from the victims' response to bereavement related trauma. A pervasive wish to bring back a loved one may be supposed to increase the likelihood of hallucinating their presence (indeed, studies of bereavement have shown this to be the case for some people). The indirect route to PE via childhood fantasy almost certainly stems from the victims' concerns about the loss of control of social and domestic domains, as Irwin's larger model suggests. Physical and sexual abuse, frequently moving home, long periods of isolation due to travel or illness, may all leave a child feeling as though they are powerless to control events in their life. In the struggle to regain control, some may turn inwards to their own uniquely private inner life of fantasy friends and make believe realms, and in doing so they may permanently open doors to heightened visualisation ability.

Also, our new model shows that PB is best modelled as a consequence, rather than a cause, of PE. This is an important result, if it can be substantiated by further research on larger samples, precisely because Irwin's original model is equivocal with respect to the effects of fantasy on PB and PE, and PB and PE themselves are supposed to show reciprocal causation with one another. If, as our model suggests, PE experience is more the result of fantasy proneness, and PB is more the result of PE, then this suggests a more phenomenological explanation for PB, over a more attributional one (i.e. that one has PB and then attributes paranormality to hallucinatory experiences). In short, where PE has previously been considered the younger brother to PB, our research suggests that it is more the 'paternal cause' of PB.

Research such as we have reported is building up an interesting and, we think, insightful picture of the causes of PE and PB. The evidence we report adds strong weight in favour of the view that traumatic events in childhood are a partial cause of heightened fantasy, and that the development of heightened fantasy has consequences for PE. Researchers in this area are only just beginning to form a picture of the causes and consequences of PB and PE. We commend the kind of model building

approach that Irwin (1993) has employed. With the modelling techniques we have used to test Irwin's model, we feel that future research can begin to move away from the rather atheoretic, piecemeal, and simple correlational approach to PB and PE, towards techniques and ideas that can enable us to truly come to understand those who experience paranormal phenomena, and those who hold paranormal beliefs.

Acknowledgements—The author would like to thank Dr Ian Deary for his assistance with statistical programming, and comments on earlier drafts of this paper. This research was funded by the Economic and Social Research Council, grant award No. R00429324200.

REFERENCES

Bentler, P. (1989). *EQS structural relations program manual.* Los Angeles, CA: BMDP Statistical Software.

Council, J. R. & Edwards, P. W. (1987). Survey of traumatic childhood events. Unpublished measure. Fargo: North Dakota State University.

Hilgard, J. R. (1970). *Personality and hypnosis: A study of imaginative involvement.* Chicago, IL: University of Chicago Press.

Hilgard, J. R. (1974). Imaginative involvement: Some characteristics of the highly hypnotizable and nonhypnotizable. *International Journal of Clinical and Experimental Hypnosis, 22,* 138–156.

Irwin, H. J. (1990). Fantasy proneness and paranormal beliefs. *Psychological Reports, 66,* 655–658.

Irwin, H. J. (1991). A study of paranormal belief, psychological adjustment, and fantasy proneness. *Journal of the American Society for Psychical Research, 85,* 317–331.

Irwin, H. J. (1992). Origins and functions of paranormal belief: The role of childhood trauma and interpersonal control. *Journal of the American Society for Psychical Research, 86,* 199–208.

Irwin, H. J. (1993). Belief in the paranormal: A review of the empirical literature. *Journal of the American Society for Psychical Research, 87,* 1–39.

Lynn, S. J. & Rhue, J. W. (1988). Fantasy proneness: Hypnosis, developmental antecedents, and psychopathology. *American Psychologist, 43,* 35–44.

Myers, S. A. (1983). The Wilson–Barber inventory of childhood memories and imaginings: Children's form and norms for 1337 children and adolescents. *Journal of Mental Imagery, 7,* 83–94.

Neppe, V. M. (1983). Temporal lobe symptomatology in subjective paranormal experients. *Journal of the American Society for Psychical Research, 77,* 1–29.

Rao, P. V. Krishna (1992). Fantasy proneness, reports of paranormal experiences and ESP test performance. *Journal of Indian Psychology, 10,* 27–34.

Thalbourne, M. A. & Delin, P. S. (1993). A new instrument for measuring the sheep-goat variable: Its psychometric properties and factor structure. *Journal of the Society for Psychical Research, 59,* 172–186.

Tobacyk, J. J. & Milford, G. (1983). Belief in paranormal phenomena: Assessment instrument development and implications for personality functioning. *Journal of Personality and Social Psychology, 44,* 1029–1037.

Tobacyk, J. J. (1988). A revised paranormal belief scale. Unpublished manuscript, Ruston, LA: Louisiana Technical University.

Wilson, S. C. & Barber, T. X. (1983). The fantasy prone personality: Implications for understanding imagery, hypnosis, and parapsychological phenomena. In Sheikh, A. A. (Ed.), *Imagery: Current theory, research, and application* (pp. 340–387). New York: Wiley.

The Journal of Parapsychology, Vol. 59, March 1995

HOW MANY FACTORS OF PARANORMAL BELIEF ARE THERE? A CRITIQUE OF THE PARANORMAL BELIEF SCALE

By Tony R. Lawrence

ABSTRACT: The purpose of this investigation was to assess critically the Paranormal Belief Scale (PBS) developed by Tobacyk and Milford (1983) and revised by Tobacyk (1988). This scale provides seven factors of paranormal belief: Traditional Religion, Psi Belief, Precognition, Spiritualism, Witchcraft, Extraordinary Life Forms, and Superstition. The revised scale consists of 26 items and has been widely used in research into the causes and consequences of belief in the paranormal. However, in its present form the PBS is based on less than strong methodology, and even though the factorial structure of paranormal belief is likely to be multidimensional, it is certainly not accounted for by a seven-factor simple structure, nor is it likely to be orthogonal. Reconstruction and reanalysis of the original 13-factor solution show that there are more likely to be four significant factors in the PBS. An in-depth review of the PBS subscales shows that subscale items lacked comprehensive coverage of their constructs and are often wrongly named. It also sheds light on why the factor structure may have turned out as it did. In concluding, I argue that a new PBS and a more appropriate definition of *paranormal* are required. Finally, a new approach to PBS scale development is argued for, helping the researcher to sift out the varieties of believer and therefore enhancing the quality of paranormal belief research.

The Psychology of Paranormal Belief

For about the last 20 years, psychologists have been paying ever greater attention to the possible causes and consequences of belief in the paranormal (see Irwin, 1993, for a comprehensive and balanced review of research in this area). One suspects that this interest is due largely to the persistence of these beliefs despite the great successes of science over the last 100 years. Parapsychologists have had an interest in paranormal belief almost since the experimental science of parapsychology began, being most specifically concerned with one particular consequence of paranormal belief—its effect on one's ostensible psychic ability (see Palmer 1971, 1972; Schmeidler, 1943; and, most recently,

I would like to acknowledge useful comment and criticism from Michael Thalbourne, Paul Kline, Chris McManus, Ian Deary, Harvey Irwin, Robert Morris, James Alcock, and one anonymous reviewer. My thanks are due especially to Dr. Tobacyk for providing additional details of analyses.

Lawrence, 1993). Research into the ostensible paranormal causes and consequences of belief in the paranormal is in itself important enough to justify interest in this area of investigation, but there are many other lines of research available to the behavioral scientist that justify taking the topic of belief in the paranormal very seriously. Indeed, psychological research on belief in the paranormal has important consequences for our understanding of these commonly held views and the people who hold them. But research conducted on belief in the paranormal is only as good as the devices used to measure it.

There are numerous questionnaires available to assess paranormal belief (Jones, Russell, & Nickell, 1977; Randall & Desrosiers, 1980; Scheidt, 1973; Thalbourne & Delin, 1993; Tobacyk, 1988). However, the most popular measure is Tobacyk and Milford's (1983; Tobacyk, 1988) Paranormal Belief Scale (PBS). Perhaps the PBS's greatest contribution to the area of paranormal belief measurement is its emphasis on the multidimensional nature of paranormal belief. For early researchers, belief in the paranormal was a unitary phenomenon, a generalized trait much like intelligence and unlike personality (e.g., Randall and Desrosiers, 1980, found that a single main factor of "supernaturalism" accounted for 70% of the variance in their factor analysis). But despite early enthusiasm for a unidimensional construct of belief in the paranormal, later factor analyses have tended to support the view that belief in paranormal phenomena is multifaceted (Clarke, 1991; Grimmer & White, 1990; Sobal & Emmons, 1982; Thalbourne & Delin, 1993).

Indeed, the PBS has seven dimensions of paranormal belief. These subscales are (in descending order of percentage of variance accounted for) Traditional Religious Belief, Psi Beliefs (mostly psychokinesis questions), Witchcraft, Superstition, Spiritualism, Extraordinary Life Forms, and Precognition. Since the development of the original scale, a number of changes led to a revised scale (Tobacyk, 1988). The revised PBS incorporates a 26-item, 7-point scale (as opposed to a 25-item, 5-point scale) and has changes in the Precognition and Witchcraft scales, and minor changes in the Extraordinary Life Forms scale.

The arguments expressed in this paper stem from the suspicion that seven dimensions of paranormal belief are simply too many. Although several researchers (e.g., Irwin, 1993) have raised specific concerns about the PBS, no one has as yet actually examined in depth its construction and content to see if the claims for seven dimensions are accurate. Consequently, in the remainder of this paper I present an in-depth critique of the PBS and its revised variant. In the first section I attempt to determine precisely how the factor analysis of the PBS could have yielded seven factors. Factor analysis is a complicated statistical technique, and the technical details presented in my reconstruction of

the original factor solution of the PBS may be skipped by those who are of a less technical persuasion. These readers are advised to move on to the last (summary) paragraph of the next section, and then to read on from the section "The Revised PBS."

The Development and Factor Analysis of the PBS

In constructing their paranormal belief scale, Tobacyk and Milford (1983) gave 61 items that in their judgment "would sample as wide a range of paranormal beliefs as possible" (p. 1030) to 391 students at Louisiana Tech University. Items were scored on a 5-point scale from 1 (strongly disagree) to 5 (strongly agree). These items were intercorrelated and factor analyzed by the principal-axis method. Promax oblique rotation was performed to assess for dependencies between factors but "revealed little evidence of dependence between factors, which allowed for the use of orthogonal rotations." The orthogonal transformation solution gave 13 factors, accounting for 60.7% of total variance.

Tobacyk and Milford (1983) do not state whether any statistically based means of factor selection were used to arrive at the original 13-factor solution, except that they go on to select only 7 of the 13 factors on the basis that these factors "appeared meaningfully interpretable as common factors." However, Dr. Tobacyk kindly informed me that he "used several criteria for the selection of factors concurrently: eigenvalue > 1.0, scree plot of eigenvalues, proportion of total variance accounted for, [and] conceptual meaningfulness of the factor dimensions" (personal communication, 1994).

It would be worthwhile to note at this juncture that Tobacyk's set of four criteria for selecting factors, although reasonable in an early perusal of the analysis, falls short of adequacy in the construction of the final scale. The trouble is that these four criteria, taken together, are likely to produce far too many factors. Also, although selection of factors that "appeared meaningfully interpretable" is a valid procedure (providing that one is clear about what would constitute a meaningful factor a priori), in Tobacyk and Milford's case the use of this criterion actually contradicts their explicit aim to let the factor analysis describe the structure of paranormal beliefs, rather than impose any a priori dimensions upon the data (à la Randall & Desrosiers, 1980). Indeed, it is hard to see how one can select factors that appear meaningfully interpretable without having some a priori conception of what should constitute a factor of paranormal belief. With at least four criteria for selecting factors from their analyses, Tobacyk and Milford's factor selection procedure may have assumed the characteristics of a projective test.

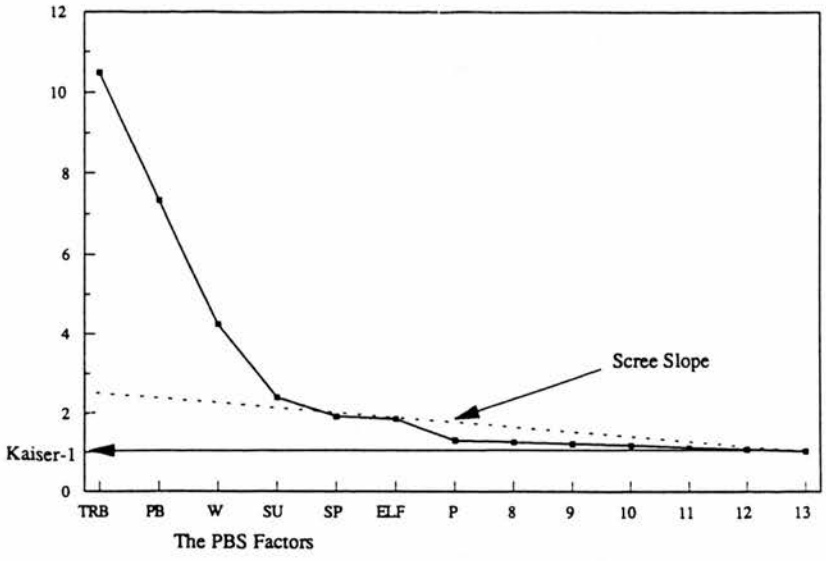

Figure 1. The 13 factors of Tobacyk and Milford's (1983) Paranormal Belief Scale.

We can find out just how many factors there would most likely be if only one, statistically reasoned factor selection procedure had been used. To start with, we can reject the simple use of the Kaiser-1 criterion on the grounds that it is too liberal (see Horn, 1965; Zwick & Velicer, 1986), that is, it gives too many factors; and indeed, in Tobacyk and Milford's analysis, it would have accounted for 13 factors alone.[1] Alternatively, looking at the proportion of variance accounted for by each variable is a redundant procedure because an eigenvalue is simply an indirect measure of proportion of variance accounted for, and thus any procedure that analyzes eigenvalues to select factors is looking at proportion of variance accounted for. Also, we cannot use the most accurate selection procedures, such as Minimum Average Partial Technique (MAP) or Parallel Analysis (PA), because these techniques require access to original data, which Tobacyk no longer has (Tobacyk, personal communication, 1994).

We are thus left with the use of a scree slope analysis which involves plotting a *visually judged* line of best fit through the smallest eigenvalues, and taking only those that come above the line. To carry out the analysis, we need to reconstruct the 13 factors of the original factor solution. We

[1]We can reasonably assume that the first 13 factors stemmed from the default application of Kaiser-1 by Tobacyk and Milford's SAS statistical package.

know the eigenvalues of the first 7 factors from data presented in Tobacyk and Milford (1983), and by making certain assumptions we can piece together the likely distribution of the original 13 factors. The assumptions that permit us to do this are as follows:

1. *Factor 7 (Precognition) is the 7th largest factor.* We are given no reason to believe that Tobacyk and Milford did not select the 7 largest factors from their 13-factor solution, and it seems unlikely that they did otherwise.
2. *Factor 13 had an eigenvalue of exactly 1.* If this was the smallest factor to make the Kaiser-1 cut, then it is most likely that it was on or very near the criterion division of 1.
3. *Factors 8 to 13 were roughly linear in distribution (i.e., triangular).* This assumption is not unreasonable. It is the weakest assumption I make, however.

We know that Tobacyk and Milford's original 13-factor solution accounted for 60.7% of the total variance in their analysis. And, by summing the percentage variance accounted for by each of their 7 selected factors, we know that these 7 factors accounted for 48.2% of the total variance. This means that, of the 13 original factors, the 6 "unknown" factors that did not make the cut accounted for 12.5% of the total variance. By invoking Assumption 1, we can deduce that Factor 8 had an eigenvalue no more than that of Factor 7 (Precognition). In this case, that is 1.281 ($2.1/100 \times 61$). So Factor 8 must be less than 1.281. From Assumption 2 we can suppose that the smallest of the 13 factors selected had an eigenvalue very close to 1. However, I shall suppose that the eigenvalue for Factor 13 was in fact 1 (in any case these calculations are unlikely to be far out). Assumption 3 may be invoked to determine the distribution of the unknown 6 factors. To do this we simply suppose the distribution of these factors to be linear and work out the gradient of the slope for these factors using basic mathematics. In this, the distance the slope must travel is 6 factors, and the height the slope must reach is from an eigenvalue of 1 to an eigenvalue of 1.281 (or 0.281 units). The gradient for this slope is given by $0.281/6 = 0.0468166$. This value is then the average distance between Factors 7 to 13. Using those first three assumptions, we can now get some idea of the distribution of Factors 8 to 13. The results of this analysis, expressed in eigenvalues, along with the known eigenvalues of the 7 factors of Tobacyk and Milford's PBS, are given in Figure 1 (n.b., Factors 8 to 13 are inferred).

Now that we have a rough plot of the eigenvalues in Tobacyk and Milford's PBS, we may try to use the more accurate and discerning scree slope criterion to select out the factors in their analysis. Using just the

scree slope analysis,[2] they would have selected Traditional Religious Belief, Psi Belief, Witchcraft, and possibly Superstition (depending on where exactly one supposes the line to be a best fit).

On these most basic analyses there are only four factors at most. The other factors might be meaningfully interpretable and have eigenvalues above 1, but they do not contribute much understanding to the concept of paranormal belief over the first three or four factors, nor do they express much more summarizing value than any one of the original 61 items. Indeed, there are conceptual reasons to suppose that the precognition factor should not be a separate factor at all, but a part of psi belief, as it surely is (see below). I shall now move on to the second major area of problems with the development of the PBS scale, where I discuss in particular concerns about the smaller factors.

The Revised PBS

In 1988 Tobacyk produced an as-yet-unpublished paper outlining changes made to the original PBS and on this basis declared a new revised PBS. In brief, the changes he made were: (a) to increase the original 25 items to 26; (b) to construct a new 7-point scale; (c) to replace the original 3-item Precognition scale with a 4-item scale of all new questions; (d) to replace 2 of 4 Witchcraft items; and (e) to replace 1 of 3 Extraordinary Life Forms items.

He argued that the changes resulted in greater reliability and validity, less restriction of range, and greater crosscultural validity in Western cultures. The inclusion of a 7-point rating scale seems fairly straightforward and makes good sense. In what follows I shall address problems with the various subscales of the PBS, before going on to consider whether Tobacyk and Milford's early evidence for construct validity is acceptable. Some of these problems are minor and can be easily remedied. Some, however, threaten the very existence of the PBS. First, I begin with the second most important factor of the PBS, the Psi subscale.

---

[2]Some earlier commentators on this paper have expressed doubts about the accuracy of the scree test over that of the Kaiser-1 criterion. However, Zwick and Velicer (1986) found it to be, on average, 35% more accurate than the Kaiser-1 criterion in selection of predetermined factors (57% vs. 22%, respectively). Earlier reports also attest to the high degree of accuracy (Cattell & Jaspers, 1967; Cliff, 1970; and Zwick & Velicer, 1982) and generally good interrater reliability among expert and naive judges (Cattell & Vogelman, 1977; Zwick & Velicer, 1982).

The Psi Subscale

Parapsychologists define *psi* as an unexplained human extrasensorimotor capability comprising extrasensory perception, or ESP (telepathy, clairvoyance, and precognition), and psychokinesis, or PK (mind over matter). Thus, as a minimum criterion for that name, any psi factor should meet the requirement that its items broadly and comprehensively deal with belief in telepathy, clairvoyance, precognition, and psychokinesis. Yet the PBS psi belief factor has only four questions, of which three address PK belief and one refers to telepathy alone. Clairvoyance is not addressed, and precognition appears on its own as the smallest factor. Plainly, the so-called Psi Belief scale is poorly named; for all intents and purposes, it is a PK belief scale. This factor of the PBS is simply too shallow to warrant either its name or the importance placed on it. Lest it be renamed the "PK belief factor," a new psi belief scale is needed.

The Precognition Subscale

Precognition is a particular kind of psi phenomenon that involves psychic perception of a future event, and along with clairvoyance and telepathy it is characterized and commonly understood to be a form of extrasensory perception. Thus, one might expect people who respond affirmatively to questions about belief in telepathy to respond affirmatively also to questions about belief in precognition, and vice versa. In short, responses to questions on precognition belief should correlate positively with those on other forms of extrasensory perception. Thus, there are sensible a priori reasons to suppose that belief in telepathy, clairvoyance, and precognition should be expressible in terms of a common factor of ESP beliefs (and to a lesser extent of PK beliefs). That the PBS showed precognition to be a separate factor from psi beliefs is perhaps the single most puzzling outcome of Tobacyk and Milford's factor analytic assessment.

It is notable that for the revised PBS (1988) Tobacyk totally replaced the three questions on precognition with four new ones (hence accounting for the increase of PBS items from 25 to 26). The reasons for doing this are outlined by Tobacyk as: (a) The original Precognition subscale has the lowest test-retest reliability of the original seven subscales (*rtt* = .60 over 4 weeks)[3]; (b) the original questions were ambiguous and could have referred to nonparanormal means of predicting the future. Tobacyk himself gives good reasons to believe that all three original

[3]Kline (1986) notes that for a minimum accepable level of reliability one should try to get at least a retest reliability of .7.

questions suffer from ambiguities of this sort. Thus, in the original PBS the ambiguity of the questions would likely have meant that a large number of respondents were not viewing these questions as being about "psychic prediction of the future," but merely about the likelihood that people could predict the future in general (say, perhaps, with meteorological techniques). Obviously these responses are not likely to covary meaningfully with the psi scale items. In fact, they are not likely to covary meaningfully with any of the subscales in the PBS. These questions are just going to sit as a separate factor on their own, which is precisely what they do.

It is clear then from Tobacyk's (1988) own paper that the Precognition subscale was far from satisfactory, both in its item content and reliability. The key issue was that the items that went into the original factor analysis related to the concept of paranormal belief only ambiguously. Tobacyk's (1988) solution to this problem was to replace these ambiguous questions with questions that are obviously related to precognition. These new questions were taken from a study (Tobacyk, Nagot, & Mitchell, 1989) that had 349 college students rate 60 beliefs about the prediction of future events. A factor analysis of these 60 items revealed seven independent dimensions of belief about the prediction of future events. According to Tobacyk (1988):

> The two clearest marker items were selected from each of the two largest paranormal precognition dimensions to form a new Precognition subscale. These dimensions and items were: Paranormal Divinatory Systems (Astrology is a way to accurately predict the future. The horoscope accurately tells a person's future) and Psychically Gifted Persons (Some psychics can accurately predict the future. Some people have an unexplained ability to predict the future). (p. 6)

Thus, from this study the four questions above were used to form the new Precognition subscale. The test-retest reliability for this scale is a reasonable .81 over 4 weeks. However, there are numerous errors and flaws with the approach to the new Precognition subscale that must render it invalid as a separate scale altogether. I shall now outline these problems with the Precognition subscale in the revised PBS.

1. *The original evidence for a factor of precognition was flawed.* As Tobacyk (1988) agrees, the original items in the PBS Precognition subscale were ambiguous. Therefore the possibility that the Precognition subscale does not actually measure belief in paranormal precognitive powers must be taken seriously. If evidence for a separate dimension of belief in precognition depends on such ambiguities, then this raises serious questions about the validity of the precognition dimension. A separate factor

of precognition cannot be upheld if the items that contributed best to that factor do not actually relate in an obvious way to precognition. Evidence for a separate factor of precognition must be derived anew.

Thus, an error was made in supposing that, although the items composing the original Precognition subscale were in need of change, the actual status of the factor of precognition needed no such scrutiny. This supposes that factors are more than the sum of their parts. In fact, the existence of factors depends crucially on the quality of their components. The subscale items come first, and the factor is always a latent variable that is to be inferred.

2. *The factorial independence of the new precognition items is not established.* This criticism stems from what has previously been argued. One should really attempt to determine anew that precognition belief is a separate aspect of belief in the paranormal. Only then can the claim be made that a separate factor of belief in precognition exists, and only then can the new precognition items be placed into such a factor. Instead, in choosing to develop new PBS Precognition scale items with reference only to items that have "prediction of future events" content, Tobacyk and his colleagues lose all possibility of establishing that a factor of belief in precognition does exist, apart from other kinds of belief in paranormal phenomena. And this is precisely what is needed to justify the inclusion of a separate factor for precognition in the first place.

3. *The Tobacyk et al. (1989) factor analysis contradicts the notion of a single factor of precognition.* Having a single factor of precognition in the original PBS was surely based on the idea that those items that were in the Precognition subscale correlated highly with one another, but not with anything much outside that subscale. Tobacyk and Milford's (1983) justification and use of orthogonal rotations implies this much. Basically, the items chosen for the original Precognition subscale represented different ways of asking: "Do you believe people can predict the future paranormally?" This is reasonable if there is good reason to suppose that a single factor exists in the factor analysis.

In their development of the Belief in Prediction of Future Events scale (BPFE), Tobacyk et al. (1989) found seven independent factors of belief in the prediction of future events. These factors underwent orthogonal rotation, producing factors that shared the least amount of variance. Subsequently, Tobacyk et al. took the strongest two marker questions from the clearest two factors of the BPFE to form their new Precognition subscale. However, in doing this they destroyed wholeheartedly the notion of a single factor of precognition. This is because the four new questions from the BPFE are taken from two independent factors (Paranormal Divinatory Systems and Psychically Gifted Persons).

In fact, two of the questions in the new Precognition subscale refer to belief in astrology and not precognition. Only one of the other two questions actually refers in a clear and definite way to precognition. In short, the new "factor" of precognition, despite its improved reliability, measures two different factors altogether, only one of which assesses belief in precognition (Psychically Gifted Persons). Changes made to the Precognition subscale may have improved the scale's reliability, but only at the expense of a near total lack of validity.

The Precognition Subscale: Conclusions

The conclusion that one must draw from an understanding of this section is that there is no good evidence for a single factor of belief in precognitive powers. Had the construction of the original PBS precognition items been conducted differently, it is most likely that no separate factor of precognition would have been found. But there may be another reason why the Precognition subscale did not form a part of a wider psi scale. As noted previously, the Psi Belief subscale contained three questions on psychokinesis and only one on ESP. Although precognition belief correlates highly with ESP belief in general, it has persistently correlated less well with PK belief in previous analyses, for reasons that should be intuitive. If items relating to Psi Belief had been truly comprehensive, the precognition items would most likely have loaded onto the factor of Psi Belief—as they have in almost every other factor analytic assessment of paranormal belief to date (see Clarke, 1991; Grimmer & White, 1990; Sobal & Emmons, 1982; Thalbourne, 1981; and, most recently, Thalbourne & Delin, 1993; but see also Thalbourne & Haraldsson, 1980).

The Witchcraft Subscale

The Witchcraft subscale of the revised PBS consists of four statements obviously relating to witchcraft and black magic. The original Witchcraft subscale had been changed because two of the original items ("Voodoo is a real method to use paranormal powers"; "There are actual cases of voodoo death") were often not familiar to foreign respondents (specifically, Finnish, German, and Polish). These items were replaced by the statements (a) "There are actual cases of witchcraft," and (b) "Through the use of formulas and incantations, it is possible to cast spells on persons."

Of these two replacement items, the first is basically good, even if a little equivocal (one could support the assertion that there are actual cases of witchcraft and still suppose that magic does not really exist). The

second statement is excellent and taps into the factor of Witchcraft without actually mentioning it explicitly, which is always a good way of constructing one's items. However, of the remaining two statements in the revised Witchcraft subscale ("Black magic really exists"; "Witches do exist") it seems clear that the statement "Witches do exist" is obviously true whether or not they really have magical powers, and the statement is likely to be endorsed as such by skeptic and Wicca worshipper alike. Perhaps this statement could be changed to something such as "Witches possess genuine magical powers." For the most part, however, there are no particular problems of major significance with the Witchcraft subscale.

The Extraordinary Life Forms Subscale

This subscale consists of three statements: (a) The abominable snowman of Tibet exists; (b) The Loch Ness monster of Scotland exists; and, (c) There is life on other planets. The items in this scale are simply of arguable paranormality. The mystery surrounding the existence (or not) of the Yeti and Nessie stems more from their apparent elusiveness than it does from any intrinsically inexplicable process of nature. If these creatures exist, they do so as a matter of natural historical fact and are as unlikely to herald a new principle of science as they are likely to warrant discarding an old one. However, the third statement in the Extraordinary Life Forms factor is probably quite useless for another reason altogether. Anyone—even a Royal Society astronomer—could agree with the statement that there is life on other planets to some degree. The real question that taps into the Extraordinary Life Forms dimension should not have been whether there was life on other planets but whether life on other planets was regularly making trips to our planet! I suspect that very few astronomers would agree with that contention. So, at the very least, the factor of Extraordinary Life Forms, which accounts for the second least percentage of variance in Tobacyk and Milford's original factor analysis, rests upon shaky foundations, theoretically and methodologically.

The Subscales of the PBS

If one wishes to develop a psychometric measure of anything of import, it is usually true that one needs a large number of items to sample the particular latent variable of interest, be it belief in the paranormal or extraversion. One needs an even larger number of items to begin to narrow down the best items to go into the refined measure. A scale that

needs only 26 items to sample no less than seven independent latent variables is probably not valid.[4] Indeed, we have seen that the Psi Belief subscale suffers from a lack of comprehensive coverage of the core psi phenomena, showing instead a prevalence of items relating solely to psychokinesis. If four or five items can get to grips with the complex attitudes and feelings underlying belief in the paranormal, then either belief in the paranormal is trivially simple or our approach to item selection is wrong.

Problems of Subscale Independence (Orthogonality)

In this section I draw on evidence from the old and new versions of the PBS. In the original paper on the old PBS, Tobacyk and Milford (1983) presented a matrix of intercorrelations between the subscales of the PBS that shed light upon the question of whether the subscales are truly uncorrelated. Clearly, examination of this matrix reveals some correlations large enough to warrant concern over subscale orthogonality. For example, Psi Belief correlates .49 with Spiritualism, .40 with Precognition (where orthogonality has already been questioned), and .34 with Witchcraft. Tobacyk and Milford (1983) dismiss these larger correlations, not by dealing with them directly, but by noting that "because these correlations are based on a 424-subject sample, statistical significance alone is not an appropriate procedure to determine relatedness between subscales. . . . The amount of shared variance (r^2) between scales appears to be a more appropriate index of relatedness" (p. 1034).

Certainly even the most minor correlations are likely to be significant with such a large sample size, and r^2 is a useful estimate of the import of the relation between two variables; but one needs to be wary indeed in throwing out evidence for relatedness (obliqueness) of subscales when the subscale correlations approach .40 and above. This wariness is especially advised given that subscale intercorrelations can only be as high as subscale reliabilities will allow. Indeed, it is a general psychometric principle that a test cannot correlate with another separate test any higher than it can correlate with itself—i.e., estimated subscale intercorrelations will always be lower than the true value for the subscale intercorrelation, because of what psychometrists call "attenuation due to unreliability." Given that generally the test-retest reliabilities for Tobacyk and Milford's original subscales fell somewhat short of being satisfactory (only Psi Beliefs and Extraordinary Life Forms were above Kline's [1993] accepted minimum of .80), we should note a particular degree of

[4]This is a point well made to me in correspondence from Paul Kline, Professor of Psychometrics of the University of Exeter, England.

caution in dismissing evidence for relatedness in the subscales. In fact, if one calculates corrected Pearson correlations (using the formula for correction of attenuation due to unreliability given by Kline, 1993, p. 14), then we see, for example, that the revised estimate of intercorrelation between Psi Beliefs and Spiritualism lies at about the .66 level, and that for Psi Beliefs and Precognition it is around .56. These figures certainly cannot be dismissed lightly, though I would advise caution in placing too strong a faith on an estimate of an estimate! Nevertheless, the original correlations are sufficient evidence for relatedness between factors for most psychometricians.

Perhaps the most revealing evidence for lack of orthogonality in the revised PBS is the fact that Tobacyk (1991) reports significant correlations between the two largest factors of his Prediction of Future Events scale and the Superstition subscale of his PBS. As the reader will be aware, Tobacyk (1988) used the two largest factors of his Prediction of Future Events scale to form the new questions for his PBS Precognition subscale. Clearly, we must regard the orthogonality of these two subscales with some suspicion. Unfortunately, Tobacyk (1988) provides no new evidence to suggest that the revised PBS contains truly orthogonal factors. Presumably, orthogonality was assumed from the previous analysis of the old PBS. But as the discussion in the first half of this section shows, the evidence for orthogonality in that scale was based on a too premature dismissal of the (in some cases quite strong) evidence for correlations between subscales.

Evidence for Construct Validity Assessed

Much of the above discussion of the subscales can be regarded as an in-depth examination of the general validity of the PBS subscales, and we have seen that there is cause to challenge both the face validity of the subscales (e.g., the item "There is life on other planets" does not even look like it taps into UFO-type beliefs) and the construct validity (e.g., both Psi Beliefs and Precognition are poor measures of their respective constructs, for reasons of comprehensive coverage and item content respectively). However, Tobacyk and Milford (1983) devote a substantial portion of their paper to demonstrating the construct validity of the PBS and its subscales empirically. Construct validity was first introduced into the psychometric literature by Cronbach and Meehl (1955) in an attempt to get around the difficulties in establishing certain specific types of validity for certain sorts of test. Construct validity refers to whether the concept or construct being measured by the scale is in reality what the scale measures. Evidence for construct validity may subsume evidence

for all other kinds of validity (concurrent, predictive, face, incremental, and differential) and may be regarded as the most general sort of validity that one can demonstrate. Thus, to show evidence for construct validity one needs to show generally that the empirical evidence derived from use of one's scale accords well with the definition of the psychological trait one has proposed.

Tobacyk and Milford (1983) established the construct validity of their scale by collecting together seven other measures of psychological traits: the Internal-External Locus of Control Scale (Rotter, 1966); the Sensation Seeking Scale (Zuckerman, 1971); the Threat Index (Hays, 1974); a self-esteem measure (Hays, 1974); the Uncritical Inferences Test (Haney, 1954); the Dogmatism Scale (Rokeach, 1960); and an Irrational Beliefs Questionnaire (Newmark, Frerking, Cook, & Newmark, 1973). They went on to correlate these measures with the global score and subscales for paranormal belief on the grounds that, "due to theoretical considerations, they were hypothesized to be related to paranormal beliefs" (p. 1034).

Tobacyk and Milford report correlations for these variables with the PBS ranging from 0-34, and they argue that these results broadly support the construct validity of the PBS. I could at this point discuss in detail the discussions of the evidence for construct validity given by Tobacyk and Milford, but I shall not do this for a simple and basic reason: None of the evidence reported by Tobacyk and Milford for construct validity is in fact evidence for such validity. Instead, I want to explain why such evidence cannot *in principle* be evidence for the construct validity of the PBS.

To establish the construct validity of a paranormal belief scale, one needs to show that it is feasible to define a construct—"paranormal beliefs"—and that the behavior of persons (both in surveys and in life generally) is consonant with such a construct. To show construct validity for the PBS, one needs to establish some basic facts. First, one needs to show concurrent validity, that is, that this measure of paranormal belief correlates reasonably well with previous measures of paranormal belief (assuming, of course, that these previous measures are valid). Second, one needs to show that the PBS can account for differences in persons that are essential and intrinsic to the notion of paranormal belief. Thus, it must be essential to any concept of paranormal belief that to have such beliefs one must have had some acquaintance with the topic of the paranormal, from whatever source (experience being the most obvious). If the PBS could not account for differences in level of acquaintance with the paranormal (which could take the form of differential or predictive validity), then we would need to question seriously the validity of the scale. For example, if we measure the degree of paranormal experience one has had, we would expect people with high experience

to show high paranormal belief, and vice versa. Any evidence to the contrary would suggest that the scale was not measuring paranormal beliefs. We could provide further evidence of this type for such things as reading books on the paranormal, visiting mediums, psychics, or astrologers for advice, and taking seriously such advice as given. Even more basically, we would need to predict, for example, that measures of paranormal belief taken now will continue to predict levels of paranormal belief in the future, and that our measure of paranormal belief does not correlate strongly with measures that are not intrinsic to paranormal belief (i.e., they are not measures of paranormal beliefs). Thus, we would not expect a measure of paranormal beliefs to correlate strongly with a measure of attitudes toward normal things (e.g., smoking, party politics, tax increases).

It should now be obvious that the "evidence" for construct validity offered by Tobacyk and Milford is simply not evidence for validity at all. Tobacyk and Milford's analyses *presume* the validity of the PBS. None of the scales used relate essentially or intrinsically to the concept of paranormal beliefs (or to events associated with paranormal phenomena). Indeed, the scales that are used to support the validity of the PBS can be used only if it is a defining characteristic of such beliefs (i.e., it is a part of the construct) that they are intrinsically dogmatic, uncritical, irrational, sensation seeking, death-threat oriented, governed by a concern with external determinants of behavior, and bound up with self esteem. In general parlance and in Tobacyk and Milford's definition of the paranormal, none of these characteristics appear as intrinsic qualities, and this is clear evidence that they cannot bear on the evidence for the construct validity of the PBS. Of course, they may relate to paranormal beliefs for reasons that are of theoretical interest, but to establish such relations requires that one has a valid measure in the first place. Thus, Tobacyk and Milford (1983) provide no good evidence that the PBS measures what it claims to measure—even if we accept some evidence for face validity of some of the items—and I say this even though I do not doubt that had such evidence been looked for, it would have been found to some extent.

To sum up this section, we may reasonably conclude that despite some evidence for face validity and content validity, the actual evidence for the validity of the PBS is very poor. Subsequent evidence from other researchers has shown some concurrent validity for the PBS—that is, it correlates with other measures of paranormal belief (Thalbourne has several unpublished studies reporting high correlations between the PBS and his sheep-goat scale e.g., Thalbourne, 1994; Thalbourne, Dunbar, & Delin, 1994; see also Irwin, 1985)—but this type of validity on its own is not sufficient to recommend the use of the scale.

Problems of Poor Scale Use

At times researchers have not used the scale in the spirit for which it was designed. The revised PBS presents seven independent subscales of paranormal belief for the researcher to use as such. The primary rationale in constructing the scale was that paranormal belief is multifaceted and not just amenable to analysis in terms of a general factor of Paranormal Belief. (Scales that have treated belief in the paranormal as an all-or-none property have tended to seem too simple and superficial; see, for example, Randall & Desrosiers, 1980, and Scheidt, 1973.) Nevertheless, in an important and highly consequential area such as the study of reasoning and belief in the paranormal, we can observe Wierzbicki (1985) correlating subjects' scores on a syllogistic reasoning task with the global score for paranormal belief and nothing else (see Irwin, 1991, for a credible attempt to approach the same issue, and Irwin, 1992, correlating childhood trauma solely with the full-scale PBS score).

There is no particular problem with reporting full-scale score correlations from a multifactorial scale (this is often done) when the meaning of the full-scale score is clear. However, I would argue that the meaning that one should give to a full-scale score on the PBS is not as clear as it might seem. The full-scale score may be tapping into general paranormal belief. But let us consider the sort of person who might score high on general paranormal belief as assessed by the PBS. This person would necessarily have expressed strong belief in so many diverse phenomena, varying widely in exotic appeal, that one might just as easily be tapping into mental health and gullibility factors too—especially since the PBS was designed to assess as wide a range of phenomena as possible. But the real problem here stems from the expansive definition of the paranormal. This leads us directly to a discussion of the definition of *paranormal*. Indeed, in this discussion we may get to the heart of the problems concerning the construct of "paranormal beliefs."

Problems in Defining "Paranormal"

Finally, I would like to raise concerns about Tobacyk's (1988) definition of the paranormal. In the revised PBS, Tobacyk used only Broad's (1949/1978) definition of *paranormality*: "explicability achieved only by major revision of the basic limiting principles of science." The problem with this definition is that it does not seem to fit the scale items adequately. Thus, with respect to Big Foot, Nessie, or extraterrestrials, surely the existence (or not) of any of these things is merely a matter of natural historical fact. Only detection of them with certainty will show that they

do exist, but their paranormal status would have to be much the same as that of the Coelacanth fish discovered off the coast of southern Africa in the 1930s. Until its discovery, the existence of the fish was not in principle ruled out, but it was just considered unlikely. But, as a matter of natural historical fact, it still existed and was later discovered.

However, Tobacyk and Milford (1983) do acknowledge that their definition of *paranormal* is problematic. Indeed, Tobacyk (1988) notes (p. 7) that the factor "Extraordinary Life Forms . . . does not strictly conform to the 'basic limiting principles' definition of paranormality." Broad's definition of *paranormal* in terms of "basic limiting principles" is not as clear-cut as it might seem. As Braude (1979) points out, "Broad has failed to explain what, in general, a phenomenon must conflict with in order to conflict with a B[asic] L[imiting] P[rinciple]" (p. 251). Clearly, a better definition is needed.

I propose that we use the far more restrictive definition of *paranormal* offered by Irwin (1993; he mistakenly attributes it to Thalbourne, 1982). Irwin defines a thing as *paranormal* if it "refers to hypothesized processes that in principle are physically impossible or outside the realm of *human capabilities* as presently conceived by conventional scientists" (p. 1; emphasis added). With this definition, Witchcraft, Psi Powers, Precognition, and Spiritualism would be retained as factors. They clearly refer to human capabilities and hypothesized processes, and all of them share a certain fundamental relationship: the dependence on unexplained principles of nature (psi processes apply to Psi, Precognition, and Spiritualism, and magical energies apply to Witchcraft).

Conclusions

From what I have outlined in this paper we can only conclude that one of psychology's most widely used scales for the assessment of belief in the paranormal, the PBS and revised PBS, is less than satisfactory for its intended purpose. I have outlined what I believe are serious objections to the view that paranormal belief can be accounted for by seven dimensions. These objections are based in part on problems with selection of questionnaire items, and for the most part they result from the all-encompassing factor selection procedure used in the original factor analysis. Quite simply, according to a more accurate factor selection criterion, there are not seven separate factors of paranormal belief, but four: Traditional Religious Belief, Psi Belief, Witchcraft, and Superstition. Further still, although there are most certainly less than seven significant factors in the PBS, those that would remain on more accurate

criteria would most likely not be orthogonal (based on a perusal of evidence reported by Tobacyk & Milford, 1983, and Tobacyk, 1991).

Clearly we need a new PBS. Much useful and insightful research has been produced with the old PBS and its revision, but the problems present in those scales necessitate a new scale, adequately designed, analyzed, and implemented. Only then can a firm foundation be made for further research in this area, and as there are no signs that research on the topic of paranormal belief (in either psychology or parapsychology) is slowing up, we can only hope to see a new PBS in the not-so-distant future.

However, before we begin to construct a new PBS, I want to note a number of concerns that may influence the exact nature of such a scale. My first concern is that the new PBS, suitably narrowed in content to contain less that is merely strange and wacky and more that is in line with the kind of humanly relevant definition of paranormal phenomena that I advocate, might run the risk of containing subscales (say of ESP, PK, and Life after Death) that contain little more than a series of bloated specifics. Cattell (1973) has shown that scales with excellent reliabilities can be obtained merely by filling each scale with items that do nothing but paraphrase each other. Thus, the two items "ESP allows people to know each other's thoughts" and "People use ESP to read other's thoughts" are little more than minor modifications on the same theme. They would correlate highly with each other and give good internal consistency and test-retest reliability, but either question alone would work just as well! My suspicion is that it is all too easy to construct face-valid and reliable belief scales using bloated specifics. This worry is compounded by a related piece of evidence, taken from the meta-analysis of the sheep-goat effect that I have previously reported (Lawrence, 1993). In that meta-analysis, Schmeidler's (1943) single-question measure of belief ("Do you believe ESP can occur under the conditions of this experiment?") gave effect-size estimates for ESP success better than almost every multi-item measure of belief in the paranormal. Clearly, some of this success can be attributed to the enhanced consistency of attitude-behavior relations found when highly specific measures of attitude are used (and it is fair to say that Schmeidler's question puts the subject on the spot). But the implication may also be that one can tap into belief just as well with one, two, or three questions as with 25 or 50.

There are two solutions to the problem of the bloated specific scale, if it exists. (We need to be sure that the problem exists by comparing, say, a six-item measure of paranormal belief—measuring clairvoyance, precognition, telepathy, PK, life after death, and magic—with a full-blown multi-item, multi-factorial scale, such as the present PBS. If they yield similar results, then bloated specifics are a problem.) First, we could use

300

the smaller, more convenient scale. Second, we could try to produce a genuine multifactorial scale that does not contain bloated specifics. One way to do this would be to index paranormal beliefs through behavioral indicators coupled with more belief-like statements, akin to the way that extraversion is measured. However, in doing this we should be aware of yet a further difficulty: To what extent can we treat paranormal belief like a basic personality trait at all? Clearly, paranormal belief is more akin to a socially held attitude, developed through paranormal experience and socialization by the wider culture. In short, although to measure someone's level of extraversion is to assess directly certain behaviors that are said to result from a basic biological/social propensity for such behavior essential to that person, to measure someone's level of paranormal belief is to measure a number of beliefs held upon evidence that is diverse socially, culturally, and psychologically.

There is a third option, however. In the third option we retain the benefits of the multifactorial scale but change the function of the scale from simply measuring beliefs to measuring beliefs in association with their most obvious and immediate social and phenomenological causes. This third option does real justice to the aim of measuring belief as socioculturally and experientially based, while actually enhancing our capacity to rationally dissect the results of pure paranormal belief research! How does one do this?

Let us consider the basic causes of paranormal belief. First and most obviously, there is personal experience of ostensibly paranormal phenomena. Second, one may be acquainted with relatives or close friends who have had paranormal experiences. Third, one may have encountered in the media popular or serious, academic or occult, or scientific or religious works relating to paranormal phenomena. Finally, one may base one's beliefs on a general world view, provided by one's education. These four basic areas then—personal, interpersonal, sociocultural, and educational—are the most basic, immediate causes of one's paranormal beliefs. Note the diversity in sources of belief. Note also that to have based one's beliefs on influences from any one of these four does not neccessarily lead to the conclusion that one has influences from any of the other three causes; in principle the four causes are independent. In practice, however, these four basic influences will be moderately related.

At present all paranormal belief researchers use measurement devices relating solely to the beliefs themselves. A most common procedure in research with these scales is to correlate paranormal belief with other psychological factors such as paranormal experience in order to better understand the mind of the paranormal believer. I would argue that this endeavor is confounded by one simple but devastating flaw—someone who scores high on a measure of paranormal belief could do

so for any combination of influences from the basic four just outlined. Thus, psychics score high on measures of paranormal belief, but so do parapsychologists and people who read about the paranormal in the *Fortean Times*. These various groups are a world apart in terms of the influence on them from the four basic causes, yet on our measure of paranormal belief and (most perniciously) in our empirical data these diverse basic differences remain undifferentiated. This obviously has disastrous, albeit hidden, consequences for the development of understanding in paranormal belief research. Thus, our basic failure to acknowledge the diversity of types of believer leads us to develop multidimensional measures of belief that are at other levels more strangely and subtly one-dimensional! This point of course applies generally to any of the present measures of paranormal belief.

Because this is a major problem, a solution becomes particularly pertinent. In short, I propose that we develop a measure more along the lines of a paranormal beliefs and influences scale (PBIS). This scale would certainly measure paranormal beliefs, but would also measure the weight of influence from the four basic causes of belief outlined above. (I realize that I am making a presumption in supposing paranormal experience to be a cause of belief, but evidence that I have reported elsewhere favors this experiential-source hypothesis over other, more attributional hypotheses; see Lawrence, Edwards, Barraclough, Church, & Hetherington, in press). Only with such a scale will we have a realistic chance of getting inside the minds of sheep and goats (whatever *species* they might be!).

Such a scale is unlikely to be produced for some time. As we have seen, building a valid and reliable measure of anything is a tricky business. One hopes that future efforts at scale construction can involve input from both parapsychologists and psychologists, both paranormalists and skeptics. This is indeed necessary if we are to break with the partisan and frankly irrational distinctions ("these are our scales and those are yours") that have split psychologists and parapsychologists with respect to the use of scales concerning paranormal belief.

Thalbourne and Delin (1993) have recently reported a revision of their Australian Sheep-Goat Scale that shows some promise in the assessment of the core concepts of paranormal belief (ESP, Psychokinesis, and Life after Death). It remains for that scale or any other to prove its worth as the scale to use. If the PBS is to be even a contender for that distinction, it will have to change. If it does change, then the enterprise undertaken in this paper will seem to this author all the less negative, and all the more productive, for it.

References

Braude, S. E. (1979). *ESP and psychokinesis: A philosophical examination.* Philadelphia: Temple University Press.

Broad, C. D. (1978). The relevance of psychical research to philosophy. In J. Ludwig (Ed.), *Philosophy and parapsychology* (pp. 43–63). Buffalo: Prometheus. (Original work published in 1949)

Cattell, R. B. (1973). *Personality and mood by questionnaire.* San Fransisco: Jossey-Bass.

Cattell, R. B., & Jaspers, J. (1967). A general plasmode for factor analytic exercises and research. *Multivariate Behavioural Research Monographs,* **3**, 1–212.

Cattell, R. B., & Vogelman, S. (1977). A comprehensive trial of the scree and KG criteria for determining the number of factors. *Multivariate Behavioural Research,* **12**, 289–325.

Clarke, D. (1991). Belief in the paranormal: A New Zealand survey. *Journal of the Society for Psychical Research,* **57**, 412–425.

Cliff, N. (1970). The relation between sample size and population characteristic vectors. *Psychometrika,* **35**, 163–178.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin,* **52**, 281–302.

Grimmer, M. R., & White, K. D. (1990). The structure of paranormal beliefs among Australian psychology students. *Journal of Psychology,* **124**, 357–370.

Haney, W. (1954). *Measurement of the ability to discriminate between inferential and descriptive statements.* Unpublished doctoral dissertation, Northwestern University.

Hays, C. (1974). *A methodological investigation of the threat index and the introduction of a short form.* Unpublished senior thesis, University of Florida.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika,* **30**, 179–185.

Irwin, H. J. (1985). A study of the measurement and correlates of paranormal belief. *Journal of the American Society for Psychical Research,* **79**, 301–326.

Irwin, H. J. (1991). The reasoning skills of paranormal believers. *Journal of Parapsychology,* **55**, 281–300.

Irwin, H. J. (1992). Origins and functions of paranormal belief: The role of childhood trauma and interpersonal control. *Journal of the American Society for Psychical Research,* **86**, 199–208.

Irwin, H. J. (1993). Belief in the paranormal: A review of the empirical literature. *Journal of the American Society for Psychical Research,* **87**, 1–39.

Jones, W. H., Russell, D. W., & Nickell, T. W. (1977). *Belief in the Paranormal Scale: An objective instrument to measure belief in magical phenomena and causes.* Psychological Documents, MS. 1577. Washington DC: American Psychological Association.

Kline, P. (1986). *A handbook of test construction.* London: Methuen.

Kline, P. (1993). *The handbook of psychological testing.* London: Routledge.

LAWRENCE, T. R. (1993). Gathering in the sheep and goats. A meta-analysis of forced choice sheep-goat ESP studies, 1947-1993. *Proceedings of Presented Papers: The Parapsychological Association 36th Annual Convention*, 75–86.

LAWRENCE, T. R., EDWARDS, C., BARRACLOUGH, N., CHURCH, S., & HETHERINGTON, F. (in press). Modelling childhood causes of paranormal belief and experience: Childhood trauma and childhood fantasy. *Personality and Individual Differences.*

NEWMARK, C., FRERKING, P., COOK, L., & NEWMARK, L. (1973). Endorsement of Ellis's irrational beliefs as a function of psychopathology. *Journal of Clinical Psychology*, **29**, 300–302.

PALMER, J. (1971). Scoring in ESP tests as a function of belief in ESP. Part I. The sheep-goat effect. *Journal of the American Society for Psychical Research*, **65**, 373–408.

PALMER, J. (1972). Scoring in ESP tests as a function of belief in ESP. Part II. Beyond the sheep-goat effect. *Journal of the American Society for Psychical Research*, **66**, 1–25.

RANDALL, T. M., & DESROSIERS, M. (1980). Measurement of supernatural belief: Sex differences and locus of control. *Journal of Personality Assessment*, **44**, 493–498.

ROKEACH, M. (1960). *The open and the closed mind.* New York: Basic Books.

ROTTER, J. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, **80**, No. 609.

SCHEIDT, R. (1973). Belief in supernatural phenomena and locus of control. *Psychological Reports*, **32**, 1159–1162.

SCHMEIDLER, G. R. (1943). Predicting good and bad scores in a clairvoyance experiment: A preliminary report. *Journal of the American Society for Psychical Research*, **37**, 103–110.

SOBAL, J., & EMMONS, C. F. (1982). Patterns of belief in religious, psychic, and other paranormal phenomena. *Zetetic Scholar*, **9**, 7–17.

THALBOURNE, M. A. (1981). Extraversion and the sheep-goat variable: A conceptual replication. *Journal of the American Society for Psychical Research*, **75**, 105–120.

THALBOURNE, M. A. (1982). *A glossary of terms used in parapsychology.* London: Heinemann.

THALBOURNE, M. A. (1994). *Further studies of the measurement and correlates of belief in the paranormal.* Manuscript submitted for publication.

THALBOURNE, M. A., & DELIN, P. S. (1993). A new instrument for measuring the sheep-goat variable: Its psychometric properties and factor structure. *Journal of the Society for Psychical Research*, **59**, 172–186.

THALBOURNE, M. A., DUNBAR, K. A., & DELIN, P. S. (1994). *An investigation into correlates of belief in the paranormal.* Unpublished manuscript, Department of Psychology, University of Adelaide, Australia.

THALBOURNE, M. A., & HARALDSSON, E. (1980). Personality characteristics of sheep and goats. *Personality and Individual Differences*, **1**, 180–185.

TOBACYK, J. J. (1988). *A revised Paranormal Belief Scale.* Unpublished manuscript, Louisiana Tech University, Ruston, LA.

TOBACYK, J. J. (1991). Superstition and belief about the prediction of future events. *Psychological Reports*, **68**, 511–512.

TOBACYK, J. J., & MILFORD, G. (1983). Belief in paranormal phenomena: Assessment instrument development and implications for personality functioning. *Journal of Personality and Social Psychology*, **44**, 1029–1037.

TOBACYK, J. J., NAGOT, E., & MITCHELL, T. (1989). Prediction of Future Events Scale: Assessment of beliefs about predicting the future. *Journal of Social Psychology*, **129**, 819–823.

WIERZBICKI, M. (1985). Reasoning errors and belief in the paranormal. *Journal of Social Psychology*, **125**, 489–494.

ZUCKERMAN, M. (1971). Dimensions of sensation seeking. *Journal of Consulting and Clinical Psychology*, **36**, 45–52.

ZWICK, W. R., & VELICER, W. F. (1982). Factors influencing four rules for determining the number of components to retain. *Multivariate Behavioural Research*, **17**, 253–269.

ZWICK, W. R., & VELICER, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, **99**, 432–442.

Department of Psychology
University of Edinburgh
7 George Square
Edinburgh
EH8 9JZ
Scotland
email: tlawrenc@afb1.ssc.ed.ac.uk

The Journal of Parapsychology, Vol. 59, June 1995

MOVING ON FROM THE PARANORMAL BELIEF SCALE: A FINAL REPLY TO TOBACYK

By Tony R. Lawrence

ABSTRACT: In this paper I reply to Tobacyk's detailed response to my critique of his Paranormal Belief Scale (see Lawrence, 1995; Tobacyk, 1995). Dealing with the points as outlined by Tobacyk, I pay particular attention to his strongest countercriticism, namely, that my use of the scree slope test is unconventional and probably invalid. Although I acknowledge that this might appear to be the case, I present new evidence to suggest that the scree slope test is almost certainly still valid and that in this and all other respects my critique remains fully viable. In concluding, I express my agreement with Tobacyk on a number of issues and end with the hope that our dialogue proves useful for the construction of a new and improved measure of paranormal belief.

Tobacyk (1995) has presented a spirited, thought-provoking, and constructive reply to my paper on the Paranormal Belief Scale (PBS) (Lawrence, 1995). His reply shows perfectly how critical dialogue can lead to significant and constructive suggestions for future research and can help to bring out problematic issues relevant to the enhancement of our understanding of complex areas of research. Also, it is clear to me from his reply that there are many areas where Tobacyk and I are in agreement. Generally, we seem to agree that we need an improvement to the PBS and that the PBS has a number of problems that only future research will be able to clear up, although we differ over just how severe these problems are. More importantly I think, both our papers have made substantive contributions on the topic of how new research in this area might proceed. In what follows I stick closely to the substance of Tobacyk's reply, dealing primarily with those areas where we disagree, before concluding my critical efforts with some final comments.

Areas Where We Disagree

Problems of Defining "Paranormal"

Although I agree with much that Tobacyk says about the problems of defining *paranormal*, Tobacyk picks up on a subtlety in Irwin's definition (the use of the word *or* rather than *and*) and goes on to show that Irwin's

definition of *paranormal* is therefore really not as restrictive as Broad's. If one sticks to the letter of Irwin's definition, then Tobacyk has a point: *Paranormal* Irwin-style may be little different from *paranormal* Broad-style. Given this subtlety, perhaps I need to give my own definition. Something is paranormal if it is "in principle physically impossible and outside the range of human capabilities as presently conceived by conventional scientists." I prefer this more restrictive definition not so much because I want it to restrict certain phenomena, but because I believe it is important for our research to be humanly relevant. I do not have space enough to go into detail, however; instead, I think the problem of defining the paranormal needs sustained critical discussion. (That will need another paper!)

We both acknowledge that the Extraordinary Life Forms subscale is not paranormal (it is, to paraphrase Tobacyk, paranormal by association). Instead, Tobacyk makes much of what I don't say about Traditional Religious Belief and Superstition. In fact, I say little about the paranormal status of traditional religious belief and superstition because my thoughts on the paranormal status of these beliefs are equivocal. I can see reasons why religious beliefs might be regarded as paranormal, but I also see reasons why calling religious beliefs paranormal might be problematic (e.g., religion depends strongly upon socially organized faith, whereas paranormal beliefs depend much more on personal human experience). As for superstition, there have been two ways of defining it in the past. First, superstition has been defined, mostly by academics, in such a way that it is identical with paranormal belief (e.g., Planer, 1988). Second, superstition has been defined, mostly in folklore, as referring more specifically to certain events and occurrences that have certain socially signified meanings (i.e., black cats are unlucky, one should throw salt over one's shoulder if one spills salt). If one adopts the first definition, then all one needs to do is decide whether the PBS is the PBS or the SBS (Superstitious Beliefs Scale). This is hardly a substantive contribution. If one adopts the second definition, then arguably superstition ceases to be paranormal. To say that black cats are unlucky is not to say that humans are possessed of some unknown capability to be unlucky in the presence of black cats. To conclude, I think that Tobacyk's reply establishes the need to discuss problems of definition as an integral part of empirical research and not as a separate component of purely philosophical or conceptual interest.

Adequacy of the PBS Factor Analysis

My discussion of the adequacy of the PBS factor analytic evidence essentially centers around three things: (1) problems of factor selection,

(2) problems of factor rotation, and (3) evidence for fewer factors from my scree slope analysis.

Problems of factor selection. On problems of factor selection, Tobacyk and I may just have to differ over Kaiser-1. He can cite studies that attest to its accuracy in some circumstances, and I can cite studies that attest to the opposite. Thus, Kline (1993) notes that

> Cattell (1978) has shown that rotation of too few factors tends to produce second order factors at the first order. Rotation of too many factors causes factors to split. A common solution to this problem and the default solution on many computer packages is to rotate all factors with eigen values or latent roots greater than one. However Cattell (1978) has shown that with large matrices this is an overestimate of the number of significant factors and Cliff (1988) has suggested that this criterion be abandoned. Barrett and Kline (1982) examined a number of methods of selecting the correct number of factors and showed that two methods appeared to reach the best solution. . . . These were the Scree Test (Cattell, 1966) and the Velicer method (Velicer, 1976). (p. 122)

At least one needs to be wary of the Kaiser-1 criterion, but given the fact that Tobacyk uses three other criteria, I think it quite fair to note caution over the stability of the factor solution. Thus, with four criteria for factor selection, does a factor have to satisfy all four criteria, or just three, or does one chop and choose criteria in order to get the factors that one desires? I think these are significant issues that may have played a role in determining the seven-factor solution.

Also, I don't think that a factor solution that accounts for 60.7% of the variance with 13 factors can be seen as satisfactory. Generally, satisfactory solutions, even when one takes into account error-ridden social and psychological data, tend to be more in the region of 60%–70% variance with 4–5 factors. Factor analysis is essentially about reducing complex sets of observed variables to their essential latent causes. When one needs 13 factors to achieve 60% accountability, or 7 factors to achieve 48% accountability (which is actually the case with the PBS), then this is hardly a victory for parsimony. Rather, it suggests that the researcher has brought together a wide and disparate set of unrelated variables, of little general explanatory consequence.

Problems of factor rotation. As my comments about the orthogonality of the PBS show, the PBS does contain a core of beliefs that do correlate—and strangely they are precisely those factors that meet the human capabilities definition of paranormal. Note that Tobacyk simply doesn't address my concerns over subscale orthogonality; instead, he simply reasserts that promax oblique rotations showed up little evidence for correlated factors. At the least I would suggest that future researchers

look strongly at the evidence for oblique factors in paranormal belief research. Consequently, I maintain that the problems of factor selection and orthogonality/rotation in the PBS remain.

Problems of the scree plot analysis. As Tobacyk points out in his reply, "Lawrence's use of the scree plot analysis to determine the meaningfulness of rotated factors is unconventional and probably invalid" (p. 6). In fact, the scree slope analysis, if performed on rotated factors, is definitely invalid. The reason for this is as follows: As Tobacyk himself notes, the scree plot analysis is used to determine the number of factors for rotation, by plotting a visually judged line of best fit through the smaller factors (expressed as eigenvalues) of the initial unrotated factor solution. Although subsequent rotation of factors does not change the overall amount of variance explained by the factor solution, it does change the specific amount of variance accounted for by each individual factor. Thus, for my scree slope to be valid, we *must* be sure we have the initial solution eigenvalues. On the face of it, I appear to have made a mistake. However, two points made to me by a colleague who specializes in psychometrics set me on a quest, the resolution of which proved highly revealing. The two "problems" pointed out to me are as follows:[1]

1. Given Tobacyk's use of principle factor analysis with Varimax rotation, the distribution of his eigenvalues is highly unlikely if they are genuinely from the rotated solution. Thus, when we look at Tobacyk's plotted eigenvalues, we see that there are some marked factors which account for a great deal of variance and many other factors which account for a very small amount. The problem is that Varimax rotation very rarely produces this kind of factor distribution. Indeed, Varimax rotations by definition tend to maximize and hence equalize the variance accounted for by factors in the rotated solution. In short, Tobacyk's data look much more as though they come from the initial solution and could only really have come from the rotated solution if Tobacyk and Milford had markedly different numbers of items for each factor.

2. Although the published data from Tobacyk and Milford (1983) are given in percentages of variance accounted for, the list of factors sent to me by Tobacyk clearly gives eigenvalues alone. But eigenvalues are of no consequence in the rotated solution, and many factor analysis packages do not give them for the rotated solution.

Now, these two pieces of evidence clearly suggest that Tobacyk has sent me the details of his *initial* solution results. We can confirm this by finding out whether Tobacyk's statistical package gave eigenvalues for the rotated solution. If the statistical package used by Tobacyk did not give eigenvalues in the rotated solution, then the validity of my analysis

[1] Thanks to Dr. Ian Deary, Professor of Psychology at the University of Edinburgh.

stands. Tobacyk used a pre-1983 version of SAS to perform his factor analysis, and so I contacted the SAS Institute in Britain to ask if that version gave eigenvalues for the rotated solution. People at the SAS Institute were very helpful, and Mark Dawson, technical support advisor, contacted SAS in the United States, where details of earlier versions were kept.

SAS in the United States still had a mainframe running a pre-1983 version of SAS(!), and so they ran a principle factor analysis with Varimax rotation on some old sample data. The results were clear: SAS pre-1983 does not give eigenvalues for the rotated solution. Thus, from the two points made above and the confirmation from SAS, we can only conclude that Tobacyk unwittingly sent me the eigenvalues for the initial solution and also apparently used them in his published report. I should perhaps add that it is not at all uncommon to find researchers reporting initial solution eigenvalues as though they were the rotated solution results. (Indeed, I gather that with SPSS the way in which data are presented often leads to this error.)

The above discovery validates my use of the scree slope analysis. However, Tobacyk argues that even if the analysis were valid, the accuracy of the slope is dubious. Thus, Tobacyk suggests that the data from Factors 14 to 25 are needed to really do the scree plot. This statement reveals two basic errors, however. First, Tobacyk forgets that he performed *principle factor analysis* and not *principle component analysis*. Thus, 13 factors is in fact how many factors need to be analyzed by the scree slope method, not 25. Second, even if principle component analysis had been used, Tobacyk is wrong in saying that there would be 25 factors. Principle component analysis extracts as many factors as variables, and in Tobacyk's original unfactored sample there were 61 variables.

In any case, Tobacyk's claim that we need those factors to identify the departure from linearity (i.e., random factors) is not appropriate. If the smaller factors were essentially random (as is usual and as Tobacyk seems to indicate), then they would generally show a linear increase in value from close to 0 to close to 1; in this case approximating this rough linearity with a smooth line would not give great inaccuracy. But even if my estimation of the plot of unknown factors is inaccurate, how much room for a radical departure from linearity is there between an eigenvalue of just above 1 (for Factor 8) and exactly 1 (for Factor 13)? It is in fact impossible for there to be a jump in the smaller factors larger than the jump that is clearly evident in the larger factors (from 2.31 for Superstition to 4.09 for Witchcraft, an increase in accountability larger almost than the whole range of accountability possible in the factors from Spiritualism to Factor 13).

Indeed, one could even argue that my placing of the line is conservative, because the rationale of the scree slope is to identify a marked increase in variance explained sufficient to show that the "mountain" of genuine factors "starts here," all the rest being mere rubble and debris. The first truly marked deviation from linearity is represented by the jump from Superstition to Witchcraft, and placing a line to reflect this gives only three factors! Thus, I think that the scree slope analysis is both valid and largely accurate (any error is unlikely to make a difference). But we could argue over small details ad infinitum, so is there any new evidence for the factorial structure of the PBS?

I have details of at least one attempt to factor analyze the PBS, performed recently by the Dutch researchers Tomic, van der Sijde, and Snel, in which they found only five factors.[2] Of Tobacyk and Milford's seven factors, only two were reproduced without change in the Dutch study: Traditional Religious Belief (their strongest factor) and Witchcraft (their third strongest). Of particular interest is that the Dutch Psi subscale consumed Tobacyk's Psi, Spiritualism, and Precognition subscales, and that the Superstition and Extraordinary Life Forms subscales loaded onto the same factor (which the Dutch group termed "Superstition"). The last factor, which the Dutch group termed "impossible," is a strange little factor composed of items that were negatively phrased. In general, then, the Dutch data, based on a moderately large sample of over 300 Dutch men and women, support fewer than seven factors. Interestingly, the Dutch Psi subscale shows a clear degree of relatedness between just those subscales of the PBS that come under the human capabilities definition of paranormal.

In short, the problems of factor selection and orthogonality remain, and on deeper analysis my scree slope analysis proves to be valid also. When we incorporate new empirical evidence bearing on the number of factors problem, we once again find fewer than seven factors.

The Correct Dimensionality of Paranormal Beliefs

In a somewhat paradoxical maneuver, Tobacyk resorts to relativistic/constructionist arguments to challenge the notion that one can determine the correct dimensionality of paranormal belief. This attempt to bypass the substance of my critique is thoroughly unsatisfactory, however. Thus, if we take the relativist position seriously, then we simply cannot legitimately ask questions about the factorial structure of

---

[2]This research is not yet published. Details about this study may be obtained from Dr. Peter van der Sijde, Transfer, Research, and Development, University of Twente, Holland.

paranormal belief at all. But Tobacyk's efforts surely presuppose that factor analysis has *some* value for research purposes.

Perhaps Tobacyk means to advocate a weak relativist position whereby cultures, people, and societies will simply differ in one or another respect on the PBS. But this weaker position turns out to express little more than trivia; we all know that people and cultures differ. The question is whether there is some invariant factor structure of beliefs within which we can express these differences. The PBS is highly problematic when viewed from this nonrelativistic perspective (e.g., how can African natives be expected to express agreement with an item on the Loch Ness monster?). I feel, in contrast to Tobacyk, that we can ask legitimate questions about the general structure of paranormal belief and that it is a challenge to devise such scales in a way that makes their underlying factors culturally invariant, even if average level of belief on the scales differs across cultures.

The Reasonableness of My Critique of the PBS

This section of Tobacyk's reply deals largely with my comment on the Extraordinary Life Forms and Precognition subscales. I believe that I have said about as much as I need to say about the Extraordinary Life Forms subscale in my original paper: (a) This subscale is not strictly paranormal (as Tobacyk agrees), and (b) it is clear that some of the items in that scale are far from adequate.

As for Tobacyk's lengthy defense of the PBS Precognition subscale, I would say that it is excellent save for one thing: Unfortunately, all the points made by Tobacyk only really bite if his Precognition subscale actually measures precognition. As I have already noted in my earlier paper, the Precognition subscale of the revised PBS barely measures precognition. First, it is composed of two factors, one of which is totally unrelated to the notion of precognitive prediction of the future, whereas the other contains only one definite item on precognition. Second, and worse still, the two factors that make up the "factor" of precognition are unrelated to each other! Consequently, anything Tobacyk says about the validity of precognition in the revised PBS needs to be weighed against the fact that his revised Precognition subscale barely measures precognition, and anything he says about the validity of the old Precognition subscale needs to be weighed against his own admission that the items in that subscale were ambiguous enough to warrant a new scale. Under these circumstances, it is highly debatable whether one can say anything of consequence about the validity of the Precognition subscale.

The Construct Validity of the PBS

Tobacyk doubts my assertion that his evidence for construct validity of the PBS is in principle not such evidence, and he asserts that in any case the PBS has "strong face and content validity" (p. 43). I do not doubt that the items on the PBS are face valid, although I am sure that some of the items on the PBS have *no* content validity (i.e., the precognition items of the new PBS, or the extraordinary life forms item). Even if the PBS did have strong face and content validity, I do not think this is sufficient to recommend its widespread use. Ultimately, construct validity is required, and I maintain that Tobacyk's initial evidence for this was no such evidence at all. Also, I disagree that construct validity is established through the construction of a rich nomological network of interrelations with other nonparanormal factors. Insofar as Tobacyk genuinely intends this to mean that the PBS interacts with other types of scales that do not themselves refer to the concept of the paranormal, such evidence could only presume construct validity and not establish it. Thus, if we establish that the PBS correlates with gullibility, as some might predict from their construct of paranormal belief, this may only establish that our PBS is an obscure indicator of gullibility.

Nevertheless, Tobacyk suggests that evidence for the construct validity of the PBS is present, and he cites one example (Tobacyk & Mitchell, 1987). I think this evidence is instructive in that it does establish some construct validity, but not of the sort that Tobacyk's construct of paranormal would require (whatever that might be). Instead, Tobacyk's evidence for construct validity supports a notion of paranormality with which we should now be quite familiar, by revealing a cluster of paranormal beliefs that all refer to "hypothesized human capabilities or processes"! Just take a look. They all fit this definition. Obliqueness rears its ugly head again. But given the remarks of Beloff (1963, p. 111), we should perhaps not be surprised at the human focus:

> Paranormal phenomena . . . almost invariably occur in connection with a human person. Now there is no a priori reason why this should be the case. One could conceive of a world in which inanimate objects took an occasional paranormal holiday during which they behaved entirely capriciously. But as it is, with the dubious exception of certain phenomena of a so-called poltergeist type and with perhaps the marginal exception of animal ESP, paranormal events are always bound up with people.

Thus, the human connection seems to be a central defining element in paranormal phenomena, of just the sort that a solid construct of the paranormal could be built on.

Final Comments

Despite our disagreement over the substance of my critique, Tobacyk makes many valuable points, including good suggestions about how future research might proceed. My own recent efforts to improve the methodological and theoretical rigor of paranormal belief research echo many of the points raised by Tobacyk (see Lawrence, Edwards, Barraclough, Church, & Heatherington, 1995). Perhaps most important, however, is the fact that Tobacyk clearly sees a greater role for understanding paranormal belief than merely enhancing knowledge. Indeed, in his parting statement emphasizing the fundamental human questions and the relevance of paranormal experience/belief research to answering these questions, Tobacyk displays a sensitivity to the wider issues sadly lacking in much contemporary parapsychology. In this respect, Tobacyk and I share the same basic attitude. It is precisely because these issues are important that we must devote sustained critical attention to the improvement of our scales for the measurement of paranormal belief (and indeed, of any variable). What is more, lest we waste time improving scales that are beyond improvement, we must focus our critical efforts on those scales that do show promise.

Tobacyk's PBS has proved a useful tool in the assessment of paranormal belief (I do not think it totally without merit!), and it is for precisely this reason that I chose to devote my attention to its shortcomings. Having performed this often painful and thankless task, I am pleased—indeed honored—to say that the sentiments expressed in the last sentence of my previous paper seem likely to be fulfilled. Even at this early stage, Tobacyk's reply *has* made my efforts seem "all the less negative, and all the more productive." So, as I say my last on this topic, I wish to thank Tobacyk for the dignified, temperate, and productive nature of his reply. May we soon see someone produce a new PBS, rising like the proverbial phoenix from the ashes of the old!

References

Beloff, J. (1963). Explaining the paranormal. *Proceedings of the Society for Psychical Research*, **42**, 101–114.

Kline, P. (1993). *The handbook of psychological testing*. London: Routledge.

Lawrence, T. R. (1995). How many factors of paranormal belief are there? A critique of the Paranormal Belief Scale. *Journal of Parapsychology*, **59**, 3–25.

Lawrence, T. R., Edwards, C., Barraclough, N., Church, S., & Heatherington, F. (1995). Modelling childhood causes of paranormal belief and

experience: Childhood trauma and childhood fantasy. *Personality and Individual Differences*, **19**, 209–215.

PLANER, F. E. (1988). *Superstition*. Buffalo, NY: Prometheus Books.

TOBACYK, J. J. (1995). What is the correct dimensionality of paranormal beliefs? A reply to Lawrence's critique of the Paranormal Belief Scale. *Journal of Parapsychology*, **59**, 27–46.

TOBACYK, J. J., & MILFORD, G. (1983). Belief in paranormal phenomena: Assessment instrument development and implications for personality functioning. *Journal of Personality and Social Psychology*, **44**, 1029–1037.

TOBACYK, J. J., & MITCHELL, T. P. (1987). The out-of-the-body experience and personality adjustment. *Journal of Nervous and Mental Disease*, **175**, 367–370.

Department of Psychology
University of Edinburgh
7 George Square
Edinburgh EH8 9JZ
Scotland
email: tlawrenc@afb1.ssc.ed.ac.uk